

Pronóstico de Demanda de Corto Plazo en Sistemas de Suministro de Energía Eléctrica utilizando Inteligencia Artificial

Tesis doctoral de

Ing. Víctor Hugo Hinojosa Mateus

aprobada por la
Facultad de Ingeniería de la
Universidad Nacional de San Juan - Argentina
para acceder al grado académico de

DOCTOR EN INGENIERÍA

Director de Tesis:

Dr. Msc. Ing. Alejandro Hoese

Instituto de Energía Eléctrica
Universidad Nacional de San Juan
Argentina

Jurado de Tesis:

Dr. Msc. Ing. Alejandro Hoese

Instituto de Energía Eléctrica
Universidad Nacional de San Juan
Argentina

Prof. Dr. François Cellier

Institute of Computational Science
ETH Zürich
Switzerland

Fecha de la defensa oral: 27 de diciembre de 2007.

Cámara Argentina del Libro – Catalogación en fuente

Hinojosa M., Víctor Hugo
Pronóstico de Demanda de Corto Plazo en Sistemas de
Suministro de Energía Eléctrica utilizando Inteligencia
Artificial.- 1ª edición – San Juan, 2008.
167 p.; 21x15 cm

ISBN 978-987-05-3838-7

1ª Edición – 22 Ejemplares

Impreso y Editado por Víctor Hinojosa, Ameghino 230 sur, San Juan, Argentina.

Impreso en Argentina, Printed in Argentine

Hecho el depósito que marca la ley 22.399

Queda prohibida la reproducción total o parcial de la obra, por cualquier medio (impreso, digital, etc.); viola derechos reservados. Cualquier utilización debe ser previamente solicitada al editor.

Todos los derechos reservados. Copyright © V. Hinojosa. 2008.

A mis hijas “Ariana y Valentina”,
a mi esposa “Viviana”,
a mis padres “V́ctor y Piedad” y
a toda mi familia...

Gracias Dios por darme fe
para seguir adelante...

PREFACIO

El desarrollo de este trabajo de investigación se realizó dentro del Programa de Doctorado en Ingeniería Eléctrica del Instituto de Energía Eléctrica de la Universidad Nacional de San Juan en la República Argentina.

La definición del tema de tesis se realizó en el Instituto de Energía Eléctrica bajo la dirección del Dr. Msc. Ing. Alejandro Hoese, quién deposito en mí toda su confianza para la realización de esta tesis doctoral. A él le hago participe de mi eterna gratitud por su permanente apoyo académico y sus valiosos consejos durante la definición, formulación y revisión del trabajo. Gracias por esas certeras correcciones y por darme la libertad necesaria para desarrollar esta investigación.

Es mí obligación expresarle mi gratitud a mi esposa Vivvy y a mis hijas Arianita y Pía Valentina por la permanente comprensión y paciencia durante la realización de la tesis. Gracias, mil gracias mis amores por las permanentes muestras de cariño y amor... sin ustedes la culminación de esta investigación no hubiera sido posible.

Por otra parte, deseo expresar mi más sincero agradecimiento al Profesor Dr. François Cellier – ETH Zürich, por su excelente disposición para aceptar el correferato de la investigación doctoral y por ayudarme con material bibliográfico de la Metodología en mi fase inicial. Siempre conté con personas de su grupo de trabajo como la Dr. Ángela Nebot y el Dr. Josep Mirats para despejar cualquier inquietud, muchas gracias.

Así mismo, hago extensivo un reconocimiento al Ing. Gabriel Argüello – Director Ejecutivo de la Corporación Centro Nacional de Control de Energía (CENACE) por haberme apoyado y ayudado en la fase final de la investigación, siempre encontré apoyo en el CENACE. Además, debo agradecer al Ing. Marco Valencia de la Dirección de Planeamiento del CENACE, por el tiempo empleado en numerosas discusiones sobre algunos temas de la tesis.

Al mismo tiempo, hago mención a todos los becarios del Instituto de Energía Eléctrica de la Facultad de Ingeniería de la Universidad Nacional de San Juan, con los que siempre pude contar.

Mi estadía en Argentina se posibilitó a través de la ayuda financiera y permanente del Servicio de Intercambio Académico – DAAD, institución a la cual le extiendo mi más eterno agradecimiento.

San Juan, 24 de diciembre de 2007

Víctor Hugo Hinojosa M.

Pronóstico de Demanda de Corto Plazo en Sistemas de Suministro de Energía Eléctrica utilizando Inteligencia Artificial

Resumen:

El Pronóstico de Demanda de Corto Plazo es una tarea de singular importancia en la Planificación, Análisis y Operación de Sistemas de Suministro de Energía Eléctrica. Los errores asociados al pronóstico de demanda pueden involucrar cuantiosas pérdidas a los Agentes del Mercado Eléctrico.

El pronóstico de demanda es fundamental para: 1) calcular los balances eléctricos y energéticos, 2) planificar la programación y ejecución de la operación, 3) elaborar los planes de expansión, inversiones y reemplazos, 4) estimar las transacciones de compra y venta de energía y servicios, 5) presupuestar los ingresos y egresos por dichas transacciones y 6) calcular el margen de pérdidas y ganancias esperado por el agente. Por esto, es muy importante que el pronóstico de demanda minimice todas las incertidumbres, considerando los cambios estructurales que se están presentando en los Mercados Eléctricos y la introducción de competencia entre los actores como resultado de la desregulación normativa.

Para el pronóstico de demanda se considera un modelo no lineal que simula cualitativa y cuantitativamente un sistema multivariable, el cual considera variables determinísticas y variables climáticas y que solo depende del horizonte a pronosticar: mediano, corto y muy corto plazo (desde 168 horas hasta 1 hora en adelante). En el corto plazo las condiciones climáticas (temperatura, humedad, velocidad del viento, visibilidad, etc.) son causas directas de la variabilidad del consumo de energía. Esta modelación cualitativa – cuantitativa tiene la ventaja de poder incluir n variables explicativas, siempre y cuando se correlacionen con la demanda sin ninguna modelación compleja, debido a las ventajas multivariables que presenta el Razonamiento Inductivo Fuzzy.

Se ha propuesto en esta tesis una nueva metodología para determinar las entradas del modelo de pronóstico, basado en Entrenamiento Supervisado. Para cumplir con este objetivo se resuelve un problema de optimización a través de Algoritmos Evolutivos considerando como función objetivo el error de simulación de un periodo histórico.

Finalmente es necesario remarcar que el concepto aquí presentado (Razonamiento Inductivo Fuzzy y el Algoritmo de Rebotes Simulados) es totalmente aplicable a cualquier problema en el que intervengan variables estocásticas y variables determinísticas.

Short Term Load Forecasting in Energy Markets using Artificial Intelligence

Abstract:

The Short Term Load Forecasting is a task of singular importance in Planning, Analysis and Operation in Energy Markets. The errors associated to load forecasting can result in considerable losses to the Agents of the Energy Market.

In fact, the load forecasting is fundamental to: 1) compute de electric and energetic balances, 2) plan the scheduling and execution of the operation, 3) elaborate the plans of expansion, investments and replacements, 4) estimate the economic/commercial transactions of energy and services, 5) budget for the revenues and outlays for the aforementioned transactions and 6) calculate the margin of expected losses and profits for the agents. For these reasons, it is very important that load forecasting minimize all uncertainties, considering the structural changes of Energy Markets and the introduction of competition between actors as a result of the deregulation.

In this research, a nonlinear model that simulates a multivariable system is presented. This model considers deterministic and climatic variables and only depends on horizon: medium, short and very short term (weekly, daily and hourly). In the short term the climatic conditions: temperature, humidity, velocity of the wind, visibility, etc. are direct causes of the variability in the consumption of energy. This Quantitative/Qualitative simulation has the advantage of including n exogenous variables whenever they are correlated with the load. This modeling is accomplished without any complex formulation due to the multivariable advantages that the Fuzzy Inductive Reasoning shows.

A new method has been proposed, using Supervised Training, to determine the inputs to the model dynamically. For this, a problem of optimization whose objective function based on error of simulation from a historic period by means of Evolutionary Algorithms is presented.

Finally, it is necessary to emphasize that the model presented (Fuzzy Inductive Reasoning and Simulated Rebounding Algorithm) is totally applicable to any problem in which stochastic and deterministic variables intervene.

ÍNDICE DE CONTENIDOS

1. MOTIVACIÓN Y FINES DEL PRONÓSTICO DE DEMANDA DE CORTO PLAZO EN SISTEMAS DE SUMINISTRO DE ENERGÍA ELÉCTRICA.....1

1.1.	INTRODUCCIÓN	1
1.2.	LA ENERGÍA Y LA DEMANDA ELÉCTRICA	1
1.2.1.	<i>Factores que afectan la Demanda Eléctrica</i>	<i>2</i>
1.2.2.	<i>Sectores de Consumo.....</i>	<i>3</i>
1.3.	PRONÓSTICO DE LA DEMANDA ELÉCTRICA	4
1.3.1.	<i>Planificación del Crecimiento</i>	<i>4</i>
1.3.2.	<i>Programación de la Operación.....</i>	<i>5</i>
1.4.	DELIMITACIÓN, ALCANCE Y OBJETIVOS DEL TRABAJO	6

2. MÉTODOS DE PRONÓSTICO DE DEMANDA USADOS EN SISTEMAS DE SUMINISTRO DE ENERGÍA ELÉCTRICA.....9

2.1.	SERIES TEMPORALES	9
2.1.1.	<i>Características de las series temporales</i>	<i>10</i>
2.1.2.	<i>Clasificación de las series temporales</i>	<i>11</i>
2.2.	INCERTIDUMBRES EN EL PDCP	13
2.2.1.	<i>Modelación general de las Incertidumbres</i>	<i>14</i>
2.2.2.	<i>Propagación de las Incertidumbres</i>	<i>14</i>
2.2.3.	<i>Incertidumbres en la medición de la demanda.....</i>	<i>15</i>
2.2.4.	<i>Incertidumbres en la obtención de datos.....</i>	<i>16</i>
2.2.5.	<i>Incertidumbres de los factores que influyen la demanda.....</i>	<i>17</i>
2.3.	MÉTODOS DE PRONÓSTICO.....	18
2.3.1.	<i>Clasificación de los métodos de pronóstico</i>	<i>18</i>
2.3.2.	<i>Métodos Estadísticos</i>	<i>19</i>
2.3.3.	<i>Métodos basados en Inteligencia Artificial.....</i>	<i>21</i>
2.3.4.	<i>Características de los métodos de referencia y FIR.....</i>	<i>22</i>
2.4.	EL SISTEMA EJEMPLO.....	23
2.4.1.	<i>Características del Sistema ejemplo</i>	<i>24</i>
2.4.2.	<i>Series temporales: demanda y temperatura</i>	<i>25</i>
2.4.3.	<i>Modelación de las Incertidumbres en el Sistema ejemplo</i>	<i>28</i>
2.4.4.	<i>Pronóstico de demanda utilizando modelos econométricos.....</i>	<i>29</i>
2.4.5.	<i>Pronóstico de demanda utilizando redes recurrentes de Jordan.....</i>	<i>30</i>

3. METODOLOGÍA DE SOLUCIÓN PROPUESTA: RAZONAMIENTO INDUCTIVO FUZZY35

3.1. RAZONAMIENTO INDUCTIVO FUZZY (FIR)	35
3.1.1. Fuzzyficación (<i>Fuzzy recoding</i>).....	36
3.1.2. Modelación cualitativa (<i>Qualitative Modeling</i>).....	38
3.1.3. Simulación cualitativa (<i>Qualitative Simulation</i>).....	39
3.1.4. Defuzzyficación (<i>Regeneration</i>)	40
3.1.5. Ejemplo de aplicación de la metodología FIR.....	40
3.2. APLICACIÓN DE LA METODOLOGÍA FIR AL PRONÓSTICO DE DEMANDA EN EL SISTEMA ECUATORIANO.....	44
3.2.1. Fuzzyficación.....	45
3.2.2. Modelación cualitativa.....	50
3.2.3. Simulación cualitativa.....	53
3.2.4. Defuzzyficación.....	55

4. COMPARACIÓN Y SELECCIÓN DE MASCARAS DE CORRELACIÓN APROPIADAS57

4.1. ESTRUCTURA DEL PROBLEMA.....	57
4.2. ESPACIO DE SOLUCIONES DEL PDCP.....	59
4.2.1. Formulación simbólica de una configuración factible (<i>CFMC</i>).....	59
4.2.2. Dimensión del problema.....	60
4.2.3. Número de configuraciones a ser evaluadas por el SRA	61
4.2.4. Función de evaluación de una configuración <i>CFMC</i>	62
4.2.5. Datos de entrenamiento y datos de validación o test	62
4.3. ALGORITMOS EVOLUTIVOS PARA LA DETERMINACIÓN DE LA CFMC ÓPTIMA	63
4.3.1. Minimizar la función objetivo mediante búsqueda local aleatoria.....	63
4.3.2. <i>Simulated Rebounding Algorithm (SRA)</i>	64

5. EVALUACIÓN DE LAS MEJORES MÁSCARAS DE CORRELACIÓN UTILIZANDO FIR & SRA69

5.1. PDCP PARA DÍAS NORMALES (NO FERIADOS)	69
5.1.1. Datos del Sistema Ejemplo.....	69
5.1.2. Corrección de valores atípicos en el PDCP para días normales.....	70
5.1.3. Armado de los conjuntos de entrenamiento y validación o test	70
5.1.4. Validez requerida de una máscara de correlación	71
5.2. PDCP UTILIZANDO FIR & SRA	71
5.2.1. Análisis de la profundidad de la máscara de correlación	71

5.2.2.	<i>Número total de configuraciones a ser evaluadas por el SRA</i>	72
5.2.3.	<i>Estructura del problema</i>	72
5.2.4.	<i>Hipótesis utilizadas en la simulación</i>	73
5.2.5.	<i>Análisis del PDCP incluyendo otras Temperaturas</i>	74
5.2.6.	<i>Análisis de sensibilidad de los k-NN</i>	75
5.2.7.	<i>Análisis de la validez de una máscara de correlación</i>	76
5.3.	ANÁLISIS DE PERFORMANCE DEL ALGORITMO SRA	77
5.4.	PDCP PARA OTRO PERIODO DE VALIDACIÓN	78
5.4.1.	<i>Base de Datos del Sistema Ejemplo</i>	78
5.4.2.	<i>Armado de los conjuntos de entrenamiento y test</i>	79
5.4.3.	<i>PDCP utilizando FIR & SRA</i>	79
5.5.	EFFECTOS ECONÓMICOS DEL PDCP EN EL DESPACHO ECONÓMICO	81
5.6.	APLICACIÓN DE LA METODOLOGÍA FIR A LA PLANIFICACIÓN DE LA OPERACIÓN EN EL SSEE DEL ECUADOR (CENACE)	89
5.6.1.	<i>Esquema del Procedimiento para el Pronóstico de Demanda</i>	89
6.	<u>RESUMEN FINAL Y CONCLUSIONES</u>	95

APÉNDICES.....**99**

A.	ALGORITMO PARA DETERMINAR EL NÚMERO DE CLASES	99
B.	ALGORITMOS PARA DETERMINAR LOS LÍMITES DE LAS FUNCIONES DE MEMBRESÍA	101
B.1.	<i>Equal Frequency Partition</i>	101
B.2.	<i>K-Means</i>	101
B.4.	<i>Fuzzy C-Means</i>	102
C:	ALGORITMO DE REBOTES SIMULADOS (SRA)	105
C.1.	<i>Algoritmos Evolutivos</i>	105
C.2.	<i>Búsqueda Local Aleatoria</i>	106
C.3.	<i>La simulación del proceso físico de rebotes inelásticos</i>	107
C.4.	<i>Estructura de vecindad del SRA</i>	109
C.5.	<i>Criterio de aceptación y detención del SRA</i>	109
C.6.	<i>Implementación práctica del SRA</i>	110
D:	RESULTADOS DE LAS SIMULACIONES	113
D.1.	<i>PDCP para la dimensión v_3 - 1 entradas para la variable demanda</i>	113
D.2.	<i>PDCP para la dimensión v_4 - 2 entradas para la variable demanda</i>	114
D.3.	<i>PDCP para la dimensión v_5 - 3 entradas para la variable demanda</i>	115
D.4.	<i>PDCP para la dimensión v_6 - 4 entradas para la variable demanda</i>	115
D.5.	<i>PDCP para la dimensión v_7 - 5 entradas para la variable demanda</i>	116
D.6.	<i>PDCP para la dimensión v_8 - 6 entradas para la variable demanda</i>	118
D.7.	<i>PDCP para la dimensión v_9 - 7 entradas para la variable demanda</i>	119

D.8.	<i>PDCP para la dimensión v_{10} - 8 entradas para la variable demanda</i>	120
E:	ANÁLISIS DEL PDCP INCLUYENDO OTRAS TEMPERATURAS	122
E.1.	<i>PDCP para la dimensión v_7</i>	122
E.2.	<i>PDCP para la dimensión v_8</i>	123
E.3.	<i>PDCP para la dimensión v_9</i>	123
E.4.	<i>PDCP para la dimensión v_{10}</i>	124
F:	ANÁLISIS DE SENSIBILIDAD DEL PARÁMETRO “k_{NN}”	125
F.1.	<i>PDCP para la dimensión v_6 - 4 entradas para la variable demanda</i>	125
F.2.	<i>PDCP para la dimensión v_7 - 5 entradas para la variable demanda</i>	125
F.3.	<i>PDCP para la dimensión v_8 - 6 entradas para la variable demanda</i>	126
F.4.	<i>PDCP para la dimensión v_9 - 7 entradas para la variable demanda</i>	126
G:	OFERTA DE GENERACIÓN DEL SISTEMA ECUATORIANO	127
H:	PLANIFICACIÓN DE LA OPERACIÓN	131
H.1.	<i>Planificación Operativa Energética</i>	131
H.2.	<i>Objetivos del Planeamiento Eléctrico</i>	133
H.3.	<i>Horizontes de Planeamiento Eléctrico de la Operación</i>	133
H.4.	<i>Programación Semanal</i>	134
H.5.	<i>Programación Diaria</i>	137
H.6.	<i>Supervisión y Control de la Operación</i>	140
REFERENCIAS BIBLIOGRÁFICAS		143
CURRICULUM VITAE		147

ÍNDICE DE FIGURAS

Figura 1.1. Curva típica de demanda diaria ($T_n = 24$ horas).....	2
Figura 1.2. Curvas de carga de diferentes sectores de consumo	3
Figura 1.3. Contexto general de actividades en los Sistemas de Energía Eléctrica	4
Figura 1.4. Etapas que componen la Programación de la Operación	5
Figura 2.1. Principales características de las series temporales	10
Figura 2.2. Función de probabilidad de la demanda.....	13
Figura 2.3. Esquema unifilar del sistema de medición.....	16
Figura 2.4. SSEE del Ecuador	24
Figura 2.5. Demanda total y temperaturas del SSEE ecuatoriano.....	25
Figura 2.6. Función de Autocorrelación de la demanda.....	26
Figura 2.7. Media y la desviación estándar de la serie de los lunes	26
Figura 2.8. Histograma de la serie de los días viernes a las 09:00 pm	27
Figura 2.9. Influencia de la temperatura en el Sistema ejemplo.....	29
Figura 2.10. Esquema del PDCP utilizando modelos SARIMA	30
Figura 2.11. Red recurrente de Jordan utilizada para el PDCP	31
Figura 2.12. Aprendizaje de la red recurrente	32
Figura 2.13. PDCP del sistema ejemplo para el mes de octubre	33
Figura 3.1. Etapas de la Metodología FIR.....	35
Figura 3.2. Fuzzyficación de la temperatura	36
Figura 3.3. Fuzzyficación del sistema lineal	41
Figura 3.4. Diagrama del proceso de Pronóstico Fuzzy	43
Figura 3.5. Señales reales y pronosticadas del Sistema lineal.....	44
Figura 3.6. Funciones de membresía para la fuzzyficación de la variable Tipo de día..	46
Figura 3.7. Funciones de membresía para la fuzzyficación de la Temperatura máxima de la Costa	47
Figura 3.8. Modelación de la Demanda del SSEE ejemplo.....	48
Figura 3.9. Funciones de membresía de “Demanda” según la variable “Tipo de Día”..	49
Figura 3.10. Fuzzyficación del sistema multivariable del SSEE ejemplo.....	51
Figura 3.11. Construcción de la matriz de comportamiento del SSEE ejemplo	53
Figura 3.12. Pronóstico fuzzy.....	54
Figura 3.13. Defuzzyficación de la demanda para el día lunes a las 19:00.....	55
Figura 5.1. Errores diarios de entrenamiento y test para v_8 y 4_NN	76
Figura 5.2. Errores diarios para el PDCP desde el 11 al 24 de septiembre de 2000	77
Figura 5.3. ACF de los residuos utilizando la Metodología FIR & SRA.....	80
Figura 5.4. Análisis del pronóstico de demanda y de sus desvíos.....	82
Figura 5.5. Características de la oferta del SSEE ecuatoriano	84
Figura 5.6. Pronóstico Semanal y Diario para la semana del 11 al 17 de septiembre de 2000	91
Figura 5.7. Pronóstico diario y horario para el día jueves 14 de septiembre de 2000....	92

Anexos

Figura C.1. Ejemplo unidimensional de rebotes simulados

Figura H.1. Programación Semanal en el MEM del Ecuador

Figura H.2. Programación Diaria en el MEM del Ecuador

ÍNDICE DE TABLAS

Tabla 2.1. Clasificación de las series temporales	11
Tabla 2.2. Índice de Correlación de las variables climáticas en Taiwán.....	17
Tabla 2.3. Precisiones de las Metodologías usadas en el SSEE de Oklahoma	18
Tabla 2.4. Modelos de Box – Jenkins o ARIMA	20
Tabla 2.5. Características de los métodos de pronóstico.....	23
Tabla 2.6. Errores del PDCP del sistema ejemplo.....	34
Tabla 3.1. Series temporales para el PDCP	44
Tabla 3.2. Fuzzyficación de la variable Tipo de día.....	45
Tabla 3.3. Fuzzyficación de la variable Temperatura máxima de la Costa.....	46
Tabla 3.4. Límites de las funciones de membresía para la modelación de la Demanda	48
Tabla 3.5. Límites máximos de las funciones de membresía para la modelación de la Demanda.....	50
Tabla 3.6. Ejemplo de una máscara de correlación	52
Tabla 4.1. Estructura de datos para la máscara de correlación (Matriz M).....	58
Tabla 4.2. Espacio de soluciones para la máscara de la Tabla 3.6	61
Tabla 4.3. Condiciones iniciales para la configuración v_8	65
Tabla 4.4. Resumen de las soluciones encontradas por el FIR & SRA para cada condición inicial	67
Tabla 5.1. Base de datos del Sistema Ejemplo.....	69
Tabla 5.2. Series temporales utilizadas en el PDCP.....	70
Tabla 5.3. Espacio de soluciones para una profundidad de 25, 49 y 73.....	72
Tabla 5.4. Resumen de los errores de entrenamiento ($E_{(t)}$) y test ($E_{TEST(t)}$) obtenidos con las distintas configuraciones de v^*	74
Tabla 5.5. Resumen de los errores obtenidos con las distintas temperaturas.....	74
Tabla 5.6. Errores obtenidos con distintos k_{NN} para distintas dimensiones de v	75
Tabla 5.7. Análisis de sensibilidad para la primera semana de test del parámetro α	78
Tabla 5.8. Segunda Base de datos de entrenamiento y test	78
Tabla 5.9. Media, desviación estándar y máximo error obtenido con distintos métodos de pronóstico	80
Tabla 5.10. Errores relativos promedios del Pronóstico de Demanda.....	81
Tabla 5.11. Escenarios de demanda considerados en el análisis.....	81
Tabla 5.12. Análisis económico para sobreestimación considerando el despacho y predespacho	86
Tabla 5.13. Análisis económico para subestimación considerando el despacho y predespacho	87
Tabla 5.14. Análisis de los sobrecostos para las Metodologías FIR & SRA y ARIMA	89
Tabla 5.15. Error relativo diario considerando el pronóstico de demanda semanal y diario.....	91
Tabla 5.16. Error relativo considerando varios horizontes horarios para el PDMCP	93

Tabla E.11. PDCP con 7 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura máxima de la Sierra

Tabla E.12. PDCP con 7 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Costa

Tabla F.1. PDCP con 4 entradas para la demanda y 4_{NN}

Tabla F.2. PDCP con 4 entradas para la demanda y 6_{NN}

Tabla F.3. PDCP con 5 entradas para la demanda y 4_{NN}

Tabla F.4. PDCP con 5 entradas para la demanda y 6_{NN}

Tabla F.5. PDCP con 6 entradas para la demanda y 4_{NN}

Tabla F.6. PDCP con 6 entradas para la demanda y 6_{NN}

Tabla F.7. PDCP con 7 entradas para la demanda y 4_{NN}

Tabla F.8. PDCP con 7 entradas para la demanda y 6_{NN}

1. MOTIVACIÓN Y FINES DEL PRONÓSTICO DE DEMANDA DE CORTO PLAZO EN SISTEMAS DE SUMINISTRO DE ENERGÍA ELÉCTRICA

1.1. INTRODUCCIÓN

A mediados de los años 1970, antes de que la eficiencia energética empiece a llamar la atención de la Opinión Pública, el pronóstico de la demanda eléctrica era un proceso relativamente simple, ya que era un reflejo directo de los requerimientos de potencia por parte de los consumidores en base a datos históricos, hoy en día es una de las tareas más importantes y de mayor responsabilidad y complejidad en Sistemas de Suministro de Energía Eléctrica (SSEE), debido a que involucran millonarias pérdidas cuando son de mala calidad. Basándose en éstos se realiza toda la Planificación y Operación de Sistemas de Potencia permitiendo evaluar las transacciones de potencia y energía, para poder así mantener un servicio continuo, confiable, seguro y económico.

Estos pronósticos son indispensables para operadores del Sistema y agentes del mercado y de ellos depende la determinación de sus metas y objetivos, siendo de fundamental importancia para:

- Calcular los balances eléctricos y energéticos.
- Planificar la programación y ejecución de la operación.
- Elaborar los planes de expansión, inversiones y reemplazos.
- Estimar las transacciones de compra y venta de energía y servicios.
- Estimar y presupuestar los ingresos y egresos por dichas transacciones.
- Calcular el margen de pérdidas y ganancias esperado por el Agente.

Por tales razones, es muy importante que el pronóstico de demanda minimice todas las incertidumbres, considerando los dramáticos cambios estructurales que se están presentando en los Mercados Eléctricos y la introducción de competencia entre actores como resultado de la desregulación normativa, donde entran en juego intereses económicos que pueden ser afectados por la ineficiencia del pronóstico, por lo que los Agentes y el propio Operador del Sistema deben trabajar con el mayor nivel de eficiencia posible.

1.2. LA ENERGÍA Y LA DEMANDA ELÉCTRICA

La demanda de potencia y el consumo de energía eléctrica son funciones no lineales en el tiempo y presentan distintos valores en los diferentes puntos geográficos de la red. Esto se debe a la naturaleza de los usuarios (residencial, comercial e industrial) y a la

intensidad y modos de uso de la potencia. Una curva característica de variación de la demanda se muestra en la Figura 1.1.

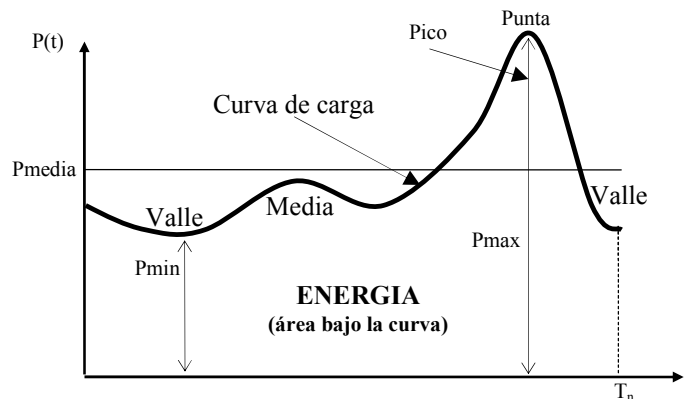


Figura 1.1. Curva típica de demanda diaria ($T_n = 24$ horas)

La demanda eléctrica es la energía consumida E_i en un periodo de tiempo T_i (ecuación (1.1)). Los valores comunes del intervalo de demanda T_i usados en la recolección de datos y facturación a usuarios son: 15, 30 ó 60 minutos.

$$E_i = \int_0^{T_i} P(t) dt \quad \Rightarrow \quad D_i = \frac{E_i}{T_i} \quad (1.1)$$

1.2.1. Factores que afectan la Demanda Eléctrica

La demanda se ve influenciada por: a) factores no controlables (sobre los cuales no se puede actuar) y b) factores controlables (voltajes, potencia, frecuencia, etc.).

a) Factores No Controlables

Costumbres de los centros de consumo: las costumbres de los diferentes centros urbanos dan como resultado variaciones de demanda diferentes. Los días de la semana también influyen en el comportamiento de los usuarios.

Factor Inercial: se produce en las mañanas de los inicios de semana y en las primeras horas después de un día feriado, resultando una demanda más baja que la media.

Estructura Económica: el consumo y la demanda de una zona eminentemente industrial son diferentes a la de una zona agrícola o de una zona turística.

Condiciones climáticas: las características de consumo de una zona varían con la temperatura, humedad, velocidad del viento, etc. siendo distintas en cada estación del año.

Factores Aislados: la presencia de un gran número de eventos, fortuitos o no, (eventos deportivos, huelgas, programas de televisión, elecciones, fallas aleatorias en el sistema de potencia, etc.) influyen el comportamiento de la demanda.

b) Factores Controlables

El más importante factor de control de la demanda es la **tarifa** como elemento regulador, ya que se puede orientar las características de consumo principalmente de los grandes usuarios (haciendo que sea más conveniente efectuar sus mayores consumos en las zonas de valle, conduciendo a un mejor factor de utilización de las unidades generadoras).

Otros factores de control de la demanda son: el voltaje, la potencia y la frecuencia, los cuales poseen un grado de incertidumbre asociado a las políticas del Ente Regulador, la operación en tiempo real por parte del Centro de Despacho, las decisiones políticas y operativas de los grandes usuarios, entre otras.

1.2.2. Sectores de Consumo

La mayoría de las empresas distribuidoras abastecen a menudo a usuarios de diferentes tipos: residenciales, comerciales, industriales y a diversos servicios que ofrece la sociedad (trenes y transporte eléctrico, alumbrado público, etc.). Cada usuario es diferente en el modo de uso de la potencia y cada sector tiende a ser semejante en el patrón de uso de la demanda [49].

Cada sector de consumo es diferente tanto en su valor medio de potencia como en el momento de ocurrencia de sus demandas máximas y mínimas. Por esto, cada Empresa de Distribución distingue el comportamiento de la carga básicamente por sectores, caracterizando a cada uno de ellos por una curva típica de carga diaria (Figura 1.2).

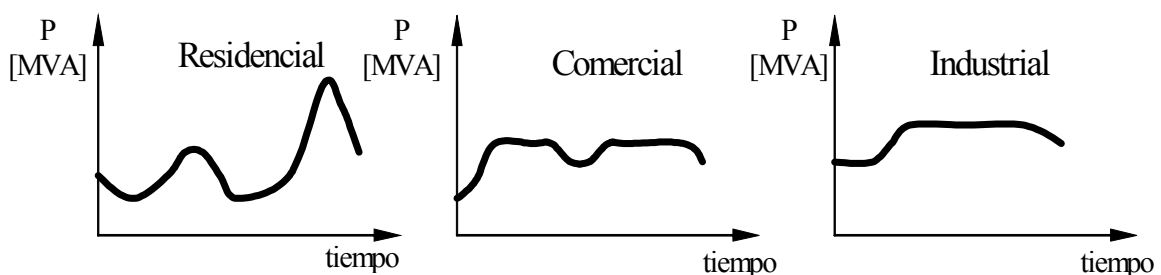


Figura 1.2. Curvas de carga de diferentes sectores de consumo

- Usuarios residenciales:** adoptan diariamente un comportamiento similar en la demanda media y pico (debido al uso de la iluminación y artefactos eléctricos). El clima es el factor de mayor peso, debido al uso de calefacción o aire acondicionado cuando la temperatura de hábitat se ha desviado de la temperatura de confort.
- Usuarios comerciales:** tienen su propia curva característica, demanda constante en la mañana y tarde asociada con el horario de comercio, con un descenso aproximadamente a las 13:00.

- c) **Usuarios industriales:** su demanda es determinada por su nivel de producción. Los servicios que se ofrecen en la sociedad se incluyen en este sector.

1.3. PRONÓSTICO DE LA DEMANDA ELÉCTRICA

El conjunto de actividades involucradas en un Sistema Eléctrico se puede sintetizar en la Figura 1.3 [16].

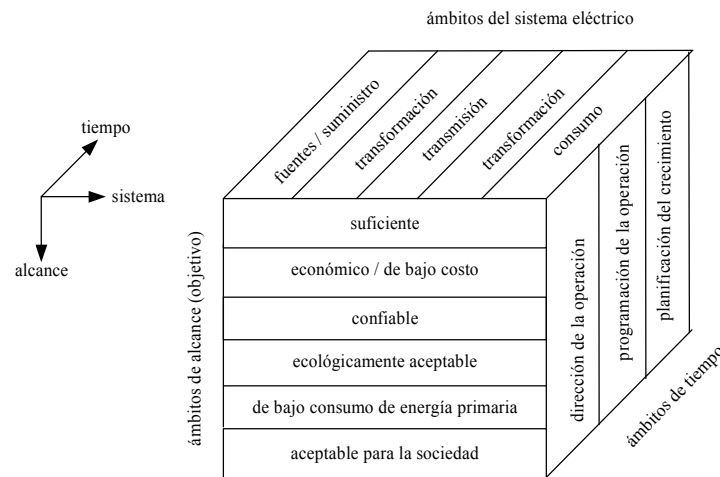


Figura 1.3. Contexto general de actividades en los Sistemas de Energía Eléctrica

El ambiente del Sistema Eléctrico se refiere a los subsistemas parciales (generación, transmisión, distribución y los consumidores). El ambiente del alcance se vincula con el objetivo de la actividad a desarrollar, que puede incorporar exigencias físicas, técnicas, económicas, medioambientales, etc. al suministro de energía eléctrica. El ambiente de tiempo se relaciona con el periodo que abarca la actividad a desarrollar y se divide en crecimiento y en operación.

Por un lado con el crecimiento de demanda se vincula la Planificación del Crecimiento o la Expansión del SSEE, que investiga la evolución futura del sistema eléctrico y que incluye periodos de planificación de hasta 30 años. Por otro lado, es tarea de la Operación del Sistema la decisión de la puesta en servicio de instalaciones de generación, transmisión y distribución a los efectos de cubrir la demanda en cada instante en forma técnica y económicamente óptima. En esta actividad se puede discernir entre la Programación de la Operación a largo, mediano y corto plazo y la Dirección de la Operación en tiempo real o Programación de muy corto plazo.

El consumo de energía, la disponibilidad de recursos fósiles, los aportes de ríos, etc. son de vital importancia en la Planificación del Crecimiento y en la Programación de la Operación [25].

1.3.1. Planificación del Crecimiento

Planificación del crecimiento de mediano y largo plazo

Comprende todos los cambios en la estructura del sistema. La potencia máxima que será requerida por los consumidores representa un indicador básico para una adecuada planificación del crecimiento del sistema. Se requiere conocer la demanda de partida y su evolución esperada en el tiempo (pronóstico de mediano y largo plazo).

Los periodos de análisis incluyen varios años hacia el futuro (2 a 20 años) y se necesita una representación de la demanda a nivel mensual, estacional y en algunos casos anual. Un aspecto importante a considerar son los escenarios alternativos dado el carácter aleatorio de la demanda.

Debe preverse tanto un sistema de generación que permita satisfacer la demanda máxima, más la reserva de potencia en cantidad suficiente para garantizar niveles mínimos de confiabilidad, como también un sistema de transmisión que posibilite el transporte de la energía desde los centros de generación hasta los consumidores al mínimo costo.

1.3.2. Programación de la Operación

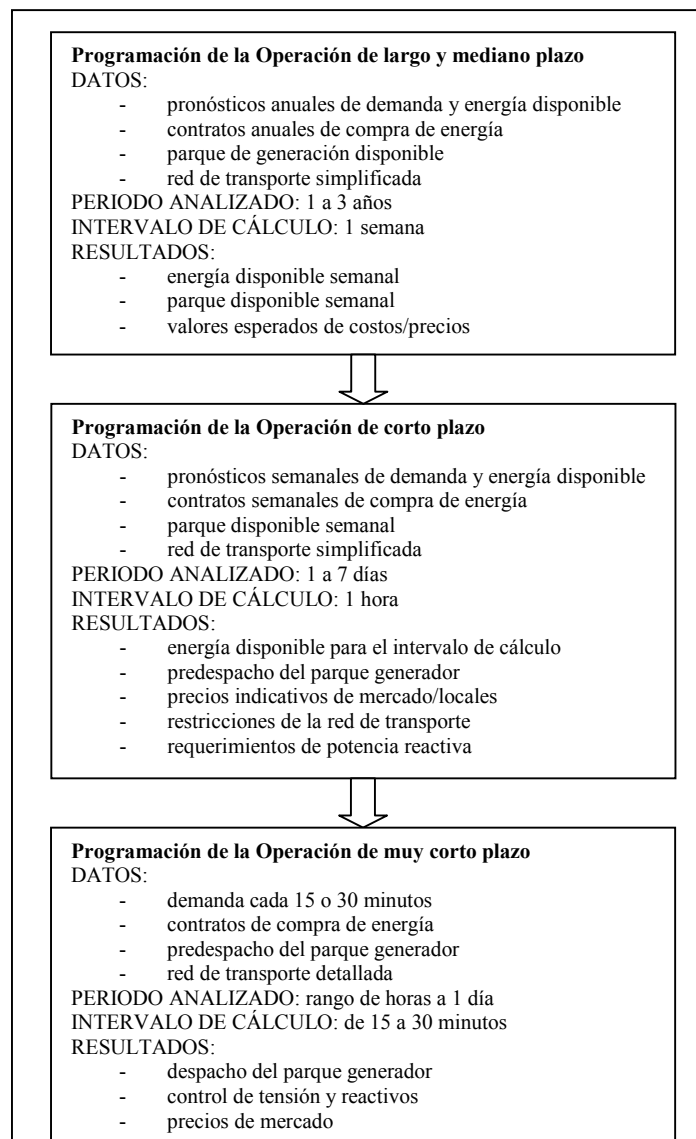


Figura 1.4. Etapas que componen la Programación de la Operación**Programación de la Operación a largo y mediano plazo**

Tiene como objetivo determinar el modo de operación de cada central, considerando las restricciones de carácter técnico, disponibilidad de recursos energéticos y de confiabilidad, de forma tal que el valor esperado de los costos, integrados en el periodo de optimización, resulte mínimo. De este proceso resulta el modo óptimo de operar los embalses de capacidad estacional (varias semanas) o mayor.

En sistemas con importantes restricciones asignadas a la red de transporte es necesario conocer la distribución geográfica de la demanda.

Programación de la Operación de corto plazo

Comprende una serie de decisiones tendientes a optimizar embalses semanales y a definir, de entre las unidades térmicas disponibles, cuáles y en qué periodo de tiempo deben ser puestas en servicio (predespacho) para suministrar con mínimo costo la energía eléctrica requerida por los usuarios. Para esto es necesario conocer los ciclos diarios y semanales de la demanda a través de curvas típicas, que representan la forma de la demanda (cronología).

Existen variables de influencia que no pueden preverse con exactitud, como los caudales de los ríos sobre los que existen centrales hidroeléctricas, la salida por falla de los bloques de generación o las variaciones en la demanda, tales variables de entrada son denominadas variables aleatorias y están dadas en general por valores pronosticados con sus respectivas probabilidades de ocurrencia.

El periodo de análisis (normalmente semanal con subperiodos horarios y diarios) depende entre otros de los ciclos de demanda, de la duración de las restricciones que afectan el servicio, del efecto “fin de semana”, del acoplamiento con la programación de largo plazo, etc.

Programación de la Operación de muy corto plazo

Tiene por objetivo calcular la reprogramación del parque hidrotérmico, con la consideración específica del control de tensiones y despacho de potencia reactiva. Se deben satisfacer además, las restricciones vinculadas al parque de generación y a la red de transporte.

La solución del problema planteado, establece el vínculo entre la programación de la operación de corto plazo (semanal) y la real del parque de generación, tanto en lo correspondiente a su potencia activa como reactiva.

1.4. DELIMITACIÓN, ALCANCE Y OBJETIVOS DEL TRABAJO

Nuestro trabajo se enfoca en el Pronóstico de Demanda de Corto Plazo, el cual sirve de base a la Programación de la Operación de corto y muy corto plazo para realizar el predespacho del parque de generación hidrotérmico. La demanda eléctrica, al ser una función no lineal de las variables intervinientes, hace que el PDCP sea básicamente un problema no lineal.

En el corto plazo las condiciones meteorológicas (temperatura, velocidad del viento, humedad, etc.) causan variación en la demanda. Además se debe incluir otras variables exógenas como: tipo de día (laboral o fin de semana), estación del año, día feriado, hora etc. para disminuir la varianza del pronóstico, ya que la demanda presenta zonas de bajo consumo (poca actividad asociada a las horas: de la noche, de un día feriado o fines de semana), zonas de consumo muy pronunciado (zonas de mucha actividad simultánea de diferentes sectores) y zonas intermedias.

El objetivo de este trabajo es analizar la aplicación de un método de modelación y simulación cualitativa – cuantitativa no lineal (Razonamiento Inductivo Fuzzy) al pronóstico de demanda de corto plazo, considerando los factores principales que afectan al pronóstico (incertidumbres) y minimizando los errores de predicción (resultados confiables).

Para esto los siguientes puntos son de primordial importancia:

- Análisis del Estado del Arte: aplicación de diferentes métodos de pronóstico utilizados. Confiabilidad de los mismos. Ventajas y desventajas.
- Horizontes de pronóstico desde 1 a 168 horas.
- Consideración de la correlación existente entre la demanda y la temperatura ambiente.
- Inclusión de variables auxiliares como: tipo de día, (laboral, feriado, etc.) a fin de disminuir la varianza del pronóstico, y reducir el espacio de posibles soluciones, los tiempos de cálculo y la complejidad del método.
- Discriminación del PDCP para cada tipo de día.

A fin de lograr este objetivo se organiza el presente trabajo de la siguiente manera:

- En el capítulo 2 se definen las series temporales y sus principales características. Se hace un análisis de las incertidumbres en el PDCP y se explican los diferentes métodos de pronóstico utilizados según el Estado del Arte de SSEE. Además, se muestran las falencias de cada metodología. Se expone de manera muy general el Razonamiento Inductivo Fuzzy (FIR) utilizado en esta investigación, detallando las ventajas y desventajas más importantes de cada metodología.

Se presenta el sistema ejemplo utilizado en este trabajo y se hace un análisis de las incertidumbres presentes. Se explica la metodología de pronóstico utilizada en el sistema ejemplo y se presenta una aplicación de Redes Neuronales Recurrentes al PDCP del sistema ejemplo.

- En el capítulo 3 se explica la metodología de solución propuesta. Se describe cada etapa de la metodología FIR y se muestra detalladamente el procedimiento y la aplicación de cada etapa al PDCP del sistema ejemplo.

Se incluye también las ventajas, desventajas y los posibles algoritmos de Inteligencia Artificial a utilizarse en cada etapa.

- En el capítulo 4 se describe cómo se formula la comparación y selección de las máscaras de correlación (Problema Combinatorial) utilizando Algoritmos Evolutivos (Simulated Rebounding), los cuales posibilitan encontrar la configuración óptima (o una cercana a la óptima) con el criterio de minimizar el error de pronóstico de una muestra.
- En el capítulo 5 se presentan los desarrollos realizados y se incluyen los resultados encontrados en la solución del problema ejemplo, para el PDCP en el SSEE Ecuatoriano.
- En el capítulo 6 se detallan las conclusiones y los aportes logrados en esta investigación.

2. MÉTODOS DE PRONÓSTICO DE DEMANDA USADOS EN SISTEMAS DE SUMINISTRO DE ENERGÍA ELÉCTRICA

Por más sofisticado que sea el método de pronóstico utilizado, el comportamiento real de un sistema siempre difiere del pronóstico hallado. En efecto, el pronóstico es simplemente la alternativa más viable que se encuentra luego de hacer una serie de razonamientos sobre futuros eventos inciertos que son importantes en la toma de decisiones presentes, basados en la información disponible y la capacidad de análisis del planificador.

Un procedimiento lógico para pronosticar el futuro es observar los datos históricos. Si la historia ha de repetirse (hipótesis de continuidad), esto es, si el comportamiento pasado puede esperarse en el futuro, es posible postular un modelo que sea representativo del proceso. Si se conoce este modelo de comportamiento (la forma exacta del modelo normalmente no puede conocerse) es factible generar pronósticos con cierto grado de exactitud. Si no se conoce el modelo de comportamiento, los datos del pasado pueden ayudar a inferir su forma.

La precisión de cualquier pronóstico de demanda está influenciada directamente por los datos disponibles, la metodología aplicada y el horizonte a planificar. Los datos de entrada deben ser:

- Confiables.
- Representativos.
- Consistentes.
- Periódicos [1].

Los datos pueden ser de corte transversal (se reúnen para un instante de tiempo, como en los censos poblacionales), o datos de series temporales.

2.1. SERIES TEMPORALES

Serie temporal es una secuencia de n observaciones (datos numéricos) ordenadas y equidistantes cronológicamente sobre una característica de una única unidad observable en diferentes instantes de tiempo, donde z_t es la observación t ($1 \leq t \leq n$) de la serie.

Definición: una serie temporal es una realización particular de una muestra procedente de un proceso estocástico univariante.

- **Proceso Estocástico (univariante):** se define como una secuencia de variables aleatorias Z_t ordenadas y equidistantes referidas a una característica de una única

unidad observable en diferentes momentos. Donde Z_t esta referida a la característica de la unidad observable considerada en el momento t .

- **Muestra:** una muestra de longitud n (finito) procedente de un proceso estocástico Z_t es un subconjunto de n componentes consecutivos.

Proceso estocástico: $Z_{-1}, Z_0, Z_1, Z_2, \dots, Z_n, Z_{n+1}, \dots$
 Muestra: Z_1, Z_2, \dots, Z_n
 Serie temporal: z_1, z_2, \dots, z_n

2.1.1. Características de las series temporales

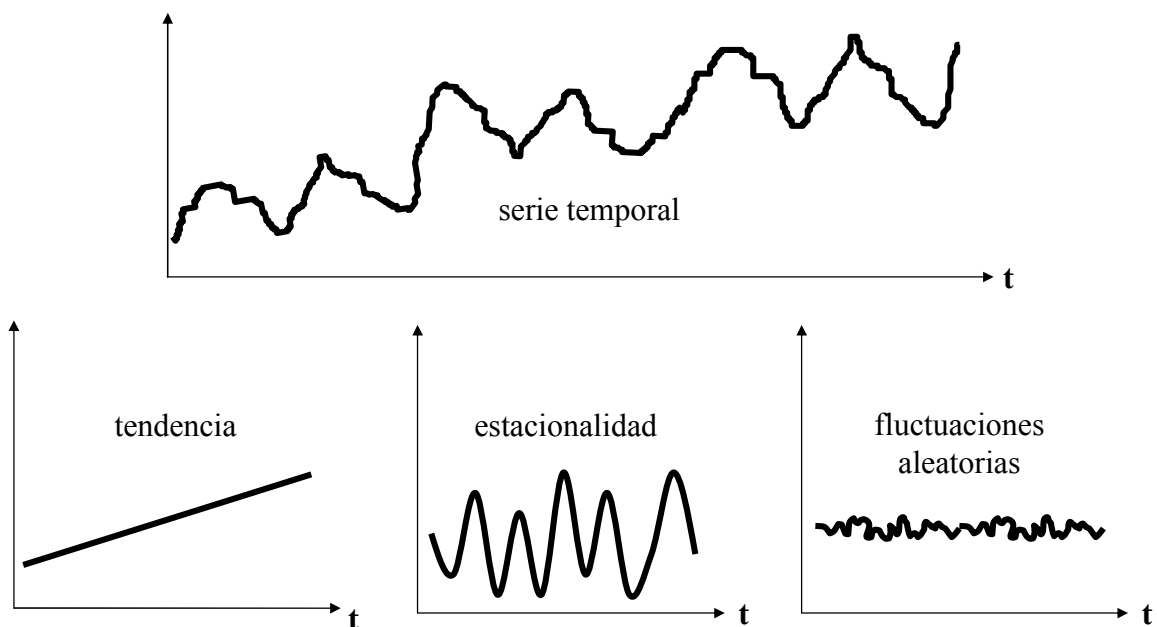


Figura 2.1. Principales características de las series temporales

Las fluctuaciones cíclicas: son curvas de largo periodo asociadas con grandes ciclos económicos. La forma de estos ciclos no es simple y en ciertas ocasiones la teoría económica no se encuentra suficientemente desarrollada como para permitir una cuantificación confiable.

La tendencia o movimientos de largo plazo de la serie: identifica los movimientos del promedio de la serie en periodos prolongados. En el caso de la demanda la tendencia se toma en periodos mensuales o anuales y son movimientos lineales o exponenciales que se relacionan con el continuo incremento en la cantidad de los usuarios y en el consumo de los mismos (crecimiento mensual o anual de la demanda).

Las fluctuaciones estacionales: son patrones que se observa con movimientos ascendentes o descendentes periódicos y que se presentan en forma regular en la serie. Al tiempo entre dos picos consecutivos se lo llama periodo estacional y puede ser: diario, semanal, trimestral, anual, etc.

Las fluctuaciones aleatorias: representan los movimientos no sistemáticos de la serie y su comportamiento es de carácter aleatorio. Aquí se acumulan todos los movimientos impredecibles de la serie que no pueden ser identificados de forma regular.

2.1.2. Clasificación de las series temporales

Natural	Artificial
Estacionaria	No estacionaria
Determinística	Estocástica
Limpia	Ruidosa
Lineal	No lineal
Continua	Discreta

Tabla 2.1. Clasificación de las series temporales

- Las series artificiales provienen de experimentos simulados libres de ruido. Las series naturales son el resultado de un determinado fenómeno físico.
- Una serie es estacionaria “estricta”, cuando la distribución de probabilidad conjunta asociada con cualesquiera n observaciones generadas por el mismo no se ve afectada por cualquier desplazamiento temporal en todos los momentos de observación [5]. Un requisito menos restrictivo, denominado estacionariedad débil de orden r , consiste en exigir que todos los momentos de órdenes r e inferiores de la estructura probabilística dependan solo de diferencias temporales. Así se habla de estacionariedad débil de segundo orden (estacionariedad en covarianza) cuando la media y la matriz de autocovarianzas existen y son acotadas. En presencia de normalidad, la estacionariedad en covarianza es equivalente a la estacionariedad estricta.
Una serie no estacionaria suele presentar ciertas tendencias marcadas (no presentan afinidad hacia algún valor constante en el tiempo) y una dispersión creciente.
- Las series pueden ser determinísticas si exhiben un comportamiento limitado de patrones, o estocásticas cuando describen un comportamiento caótico.
- Una serie ruidosa es frecuentemente producida por: a) problemas o fallas en los equipos de medición (errores sistemáticos), b) por el ruido inherente al mismo sistema de medición y c) el ruido proveniente de fuentes de perturbación exteriores al sistema. En una serie limpia los equipos no tienen ninguna fuente de error o contaminación (situación ideal).
- El sistema del cual la serie es determinada puede ser lineal o no lineal, con respecto al comportamiento autoregresivo de sus observaciones.
- La serie temporal puede ser continua o discreta.

Procesos Estocásticos Estacionarios

Si se considera n observaciones $z_{t1}, z_{t2}, \dots, z_{tn}$, generadas por un proceso estocástico (Z_t) en cualesquiera momentos distintos t_1, t_2, \dots, t_n ; el proceso es considerado estacionario si y solo si satisface las condiciones:

- a) Existe una función de densidad no singular que caracteriza la distribución de probabilidad conjunta de los n componentes del proceso asociados con las n observaciones para cualquier entero $n \geq 1$ (finito) y,
 - b) La función de probabilidad coincide con la de los n componentes del proceso asociados con otras n observaciones $z_{t1+k}, z_{t2+k}, \dots, z_{tn+k}$ para cualquier entero $k \neq 0$ (finito).
- Cuando $n = 1$, las condiciones a) y b) implican que la distribución de Z_t existe y es la misma en todo instante t . Esta distribución estará caracterizada por una función de densidad $p(x)$, donde x es una variable real, de manera que la **media** μ_Z y la **varianza** σ_Z^2 de un proceso estacionario se pueden calcular con la ecuación (2.1).

$$\begin{aligned}\mu_Z &\equiv E[Z_t] = \int_{-\infty}^{+\infty} xp(x) dx \\ \sigma_Z^2 &\equiv E[(Z_t - \mu_Z)^2] = \int_{-\infty}^{+\infty} (x - \mu_Z)^2 p(x) dx\end{aligned}\tag{2.1}$$

En un proceso estacionario estos dos momentos son constantes y miden, respectivamente, el nivel con respecto al cual evoluciona el proceso y la dispersión del mismo alrededor de dicho nivel. Dado que la distribución descrita por $p(x)$ es la misma en todo instante t , su forma puede inferirse a partir del histograma de las observaciones que conforman la serie generada por el proceso.

- Cuando $n = 2$, las condiciones a) y b) implican que la distribución conjunta de Z_t y Z_{t+k} existe para cualquier entero $k \neq 0$ y es la misma en todo instante t . Esta distribución está caracterizada por una función de densidad $p_k(x_1, x_2)$, donde x_1 y x_2 son variables reales; la forma de la función de densidad puede inferirse a partir de representaciones en el plano de pares (z_t, z_{t+k}) de observaciones de la serie generada por el proceso, separadas por un intervalo o retardo k dado. La **covarianza de orden k** entre Z_t y Z_{t+k} es la misma en todo instante t y se la calcula mediante la ecuación (2.2).

$$\begin{aligned}\Gamma_k &\equiv Cov[Z_t, Z_{t+k}] = E[(Z_t - \mu_Z)(Z_{t+k} - \mu_Z)] \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1 - \mu_Z)(x_2 - \mu_Z) p_k(x_1, x_2) dx_1 dx_2\end{aligned}\tag{2.2}$$

Tanto la forma funcional de $p_k(x_1, x_2)$ como Γ_k sólo dependen del retardo k y no del instante t . La **Autocorrelación** (simple) de orden k se define como el coeficiente de correlación lineal simple entre Z_t y Z_{t+k} y se calcula mediante la ecuación (2.3).

$$\rho_k = \frac{\text{Cov}[Z_t, Z_{t+k}]}{\sqrt{\sigma_{Z_t}} \sqrt{\sigma_{Z_{t+k}}}} \quad (2.3)$$

Como funciones del retardo k , las secuencias γ_k y ρ_k para $k \geq 0$ se denominan **función de autocovarianza** y **función de autocorrelación (ACF)**.

Como estimadores razonables de la media y la varianza de un proceso estacionario, se puede considerar la media muestral $\hat{\mu}_Z$ y la varianza muestral $\hat{\sigma}_Z^2$ (ecuación (2.4)).

$$\begin{aligned} \hat{\mu}_Z &= \frac{1}{N} \sum_{t=1}^N Z_t \\ \hat{\sigma}_Z^2 &= \frac{1}{N} \sum_{t=1}^N (Z_t - \hat{\mu}_Z)^2 \end{aligned} \quad (2.4)$$

Estos estimadores proporcionan valores numéricos \bar{z} y s^2 , cuando en la ecuación (2.4) se reemplaza cada variable aleatoria Z_t por su valor observado z_t .

2.2. INCERTIDUMBRES EN EL PDCP

La demanda eléctrica, sujeta a errores de medición, puede interpretarse como la convolución entre la función de probabilidad delta de Dirac (δ_D) en el valor medio de la demanda y la función de probabilidad del error del punto de medición, como indica esquemáticamente la Figura 2.2.

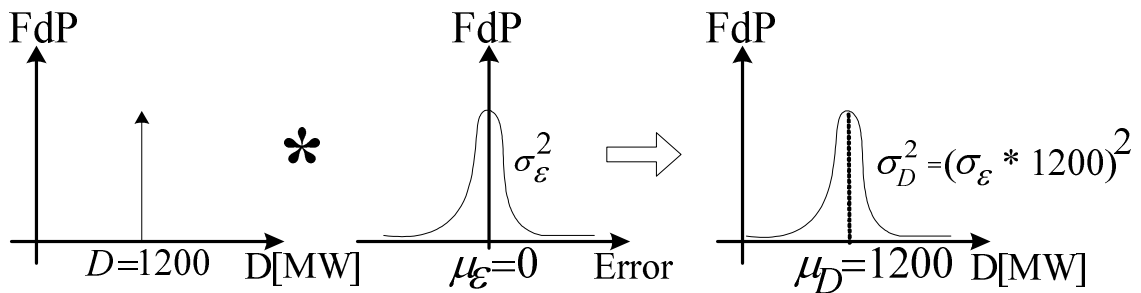


Figura 2.2. Función de probabilidad de la demanda

La función de probabilidad del error es conocida como ruido blanco en procesos estocásticos, y es representada por una distribución normal de media cero, desviación estándar σ_ϵ y autocorrelaciones nulas [32].

2.2.1. Modelación general de las Incertidumbres

Se define como Incertidumbre γ_z de una variable aleatoria z a la medida de la desviación relativa del valor real de dicha variable respecto de su valor medio (conocido) [24].

- a) Cuando la función de probabilidad de z es conocida, se puede describir la incertidumbre γ_z de z mediante la ecuación (2.5).

$$\gamma_z = \frac{S_z}{\bar{z}} \quad (2.5)$$

- b) Si la función de probabilidad de z no es conocida, entonces se debe:

- i) Determinar la máxima desviación posible Δ_z , ó,
- ii) Estimar Δ_z .

De esta manera puede describirse z por medio de una función de probabilidad triangular, donde es válido que $z_{\min} < z < z_{\max}$, entonces la incertidumbre la podemos calcular mediante la ecuación (2.6).

$$\left. \begin{array}{l} z_{\max} \leq \bar{z} + \Delta_z \\ z_{\min} \geq \bar{z} - \Delta_z \end{array} \right\} \Rightarrow \gamma_z = \frac{\Delta_z}{\bar{z}} \quad (2.6)$$

De acuerdo al modelo de incertidumbre que se aplique en un proceso de decisión, se producen los siguientes casos [18]:

- En el caso a): una situación de riesgo.
- En el caso b) inciso i): una situación adaptiva.
- En el caso b) inciso ii): una situación incierta.

2.2.2. Propagación de las Incertidumbres

Sea una función $q(x,y)$, donde las variables x e y , están sujetas a errores de medición, los cuales se propagan en la determinación de q .

Si se realizan N mediciones (x_i, y_i) , de las N mediciones de x_i es posible calcular la media \bar{x} y la desviación estándar s_x , similarmente para los y_i es posible calcular \bar{y} y s_y . Entonces, la cantidad q para cada medición i , se calcula como: $q_i = q(x_i, y_i)$ para $i=1, \dots, N$.

Asumiendo que las incertidumbres asociadas a las variables x e y son pequeñas, las mediciones x_1, \dots, x_N e y_1, \dots, y_N serán cercanas a \bar{x} e \bar{y} respectivamente, con lo cual es válida la ecuación (2.7).

$$q_i(x_i, y_i) \approx q(\bar{x}, \bar{y}) + \frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y}) \quad (2.7)$$

En la ecuación (2.7), las derivadas parciales son tomadas en el punto $x=\bar{x}$ e $y=\bar{y}$, por lo que son constantes para $i=1,\dots,N$. Con esta aproximación se puede calcular la media \bar{q} con la ecuación (2.8).

$$\bar{q} = \frac{1}{N} \sum_{i=1}^N q_i = \frac{1}{N} \sum_{i=1}^N \left[q(\bar{x}, \bar{y}) + \frac{\partial q}{\partial x} (x_i - \bar{x}) + \frac{\partial q}{\partial y} (y_i - \bar{y}) \right] \quad (2.8)$$

En la ecuación (2.8), los últimos dos términos del corchete son exactamente cero, con lo que se obtiene la ecuación (2.9).

$$\bar{q} = q(\bar{x}, \bar{y}) \quad (2.9)$$

La varianza s_q^2 puede calcularse por medio de la ecuación (2.10).

$$s_q^2 = \frac{1}{N} \sum (q_i - \bar{q})^2 = \frac{1}{N} \sum \left[\frac{\partial q}{\partial x} (x_i - \bar{x}) + \frac{\partial q}{\partial y} (y_i - \bar{y}) \right]^2 \quad (2.10)$$

$$s_q^2 = \left(\frac{\partial q}{\partial x} \right)^2 \frac{1}{N} \sum (x_i - \bar{x})^2 + \left(\frac{\partial q}{\partial y} \right)^2 \frac{1}{N} \sum (y_i - \bar{y})^2 + 2 \frac{\partial q}{\partial x} \frac{\partial q}{\partial y} \frac{1}{N} \sum (x_i - \bar{x})(y_i - \bar{y})$$

Se define ahora la covarianza s_{xy} de 2 variables x e y mediante la ecuación (2.11).

$$s_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \quad (2.11)$$

Por lo tanto, la varianza ocasionada por el error de medición de x e y es igual a la ecuación (2.12).

$$s_q^2 = \left(\frac{\partial q}{\partial x} \right)^2 s_x^2 + \left(\frac{\partial q}{\partial y} \right)^2 s_y^2 + 2 \frac{\partial q}{\partial x} \frac{\partial q}{\partial y} s_{xy} \quad (2.12)$$

Con lo cual la incertidumbre γ_q de la variable q se obtiene como propagación de γ_x , γ_y , y γ_{xy} según la ecuación (2.13) [44].

$$\gamma_q = \sqrt{\left(\frac{\partial q}{\partial x} \frac{\bar{x}}{\bar{q}} \right)^2 \gamma_x^2 + \left(\frac{\partial q}{\partial y} \frac{\bar{y}}{\bar{q}} \right)^2 \gamma_y^2 + 2 \left(\frac{\partial q}{\partial x} \frac{\bar{x}}{\bar{q}} \right) \left(\frac{\partial q}{\partial y} \frac{\bar{y}}{\bar{q}} \right) \gamma_{xy}} \quad (2.13)$$

2.2.3. Incertidumbres en la medición de la demanda

Los puntos de entrega del sistema de potencia a las empresas de distribución constan de un transformador de corriente, un transformador de tensión y un medidor de energía como se muestra en la Figura 2.3.

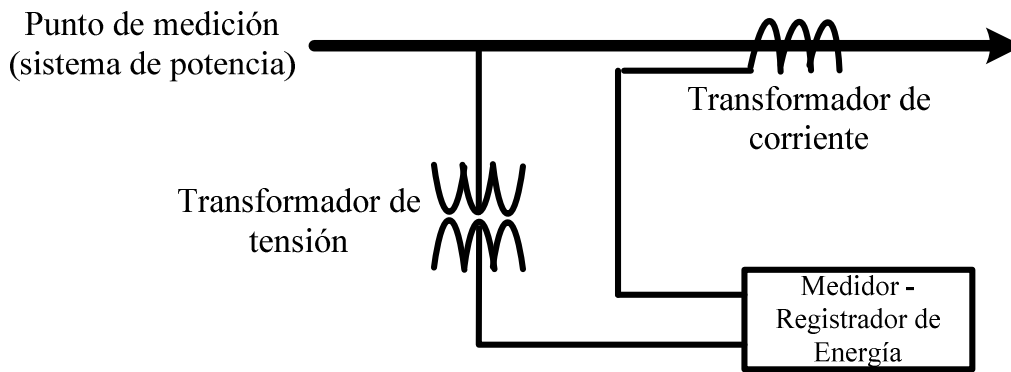


Figura 2.3. Esquema unifilar del sistema de medición

El error de la medición es igual al error introducido por el medidor de energía más el error introducido por los transformadores de corriente y tensión. Aplicando la ecuación (2.12) y considerando que los errores de medición no están correlacionados, la incertidumbre en el punto de medición viene dada por la ecuación (2.14).

$$\gamma_D = \frac{s_D}{D} = \sqrt{\left(\frac{s_I}{I}\right)^2 + \left(\frac{s_V}{V}\right)^2} + \frac{s_E}{E} = \sqrt{\gamma_I^2 + \gamma_V^2} + \gamma_E \quad (2.14)$$

donde;

- γ_j : incertidumbre del error de medición de la variable j
- s_j : desviación estándar del error de medición, donde $j = D, I, V, E$
- D : demanda de potencia [W]
- I : corriente [A]
- V : tensión [V]
- E : energía [Wh]

Se asume que la medición de la demanda no tiene ninguna otra fuente de error, es decir se desprecian los errores por transmisión de datos y los errores sistemáticos.

2.2.4. Incertidumbres en la obtención de datos

La calidad de los modelos de pronóstico depende en gran medida de los datos incluidos en las series temporales que conforman sus entradas, por lo que se debe realizar un tratamiento previo de los valores atípicos.

Las observaciones que son rechazadas se consideran como datos atípicos y no son representativos del comportamiento real de la serie en condiciones normales.

- En la referencia [12] se considera una distribución estadística T y se define un nivel de confianza para identificar los datos atípicos.
- La referencia [43] considera una distribución de datos normal, donde el 99,7% de los datos deben estar entre la media +/- tres desviaciones estándar.

Método formal de rechazo de datos – Criterio de Chauvenet [44]

Este criterio rechaza una observación de una muestra de tamaño N , si la desviación de la media es mayor que la que corresponde a una probabilidad de $1/(2N)$. La probabilidad se calcula bajo el supuesto de una distribución normal.

Para esto, se calcula la media y la desviación estándar de las mediciones y se encuentra mediante la ecuación (2.15), un valor t que se requiere analizar.

$$t_{sus} = \frac{|y_{sus} - \bar{y}|}{s_y} \quad (2.15)$$

donde; y_{sus} : valor de demanda horaria que se sospecha es atípico

Se encuentra ahora la probabilidad de t , número de mediciones desviadas que se espera (ecuación (2.16)).

$$n = N * Prob(fuera \rightarrow t_{sus} \sigma) \quad (2.16)$$

Si $n < 1/2$ entonces, este criterio rechaza el valor.

Algunos investigadores creen que los datos nunca deberían ser rechazados, un razonable compromiso es usar este criterio y una vez hecha esta identificación calcular los resultados incluyendo y excluyendo los datos atípicos para ver como se ve afectado el resultado final.

2.2.5. Incertidumbres de los factores que influncian la demanda

En este grupo se incluyen todos los factores (controlables y no controlables) que afectan la demanda eléctrica (ítem 1.2.2.).

Influencia de la correlación con variables climáticas

En la referencia [50] se estudia el sistema eléctrico de Taiwán, donde al PDCP se incorpora registros climáticos distribuidos de las ciudades de Taipei (norte), Taichung (centro) y Kaohsiung (sur). Los coeficientes de correlación (Pearson) entre la demanda y las variables climáticas se muestran en la Tabla 2.2.

Día	Variable	Verano	Invierno	Otoño	Primavera
Lunes a Viernes	Temperatura	0,8311	0,3813	0,4929	0,4459
	Humedad	-0,7694	-0,5667	-0,6078	-0,6764
	Vel. Viento	0,4612	0,2831	0,4098	0,4003
Fin de semana	Temperatura	0,4869	0,2314	0,4825	0,3742
	Humedad	-0,4224	-0,3502	-0,3608	-0,4790
	Vel. Viento	0,2822	0,0912	0,2044	0,1675

Tabla 2.2. Índice de Correlación de las variables climáticas en Taiwán

La temperatura es la variable con el más alto nivel de correlación. Se desprecian las variables climáticas cuyo coeficiente de correlación sea menor a +0,5.

En la Referencia [19], se analiza una Empresa de Distribución de Oklahoma con 450000 clientes y un pico aproximado en el verano de 1984 de 3.000 MW. La Empresa realiza el PDCP mediante un análisis histórico de la demanda, donde la experiencia del planificador juega un papel importante en el método. Los errores del PDCP se resumen en la Tabla 2.3.

Año 1984	ARIMA	ARIMAX	Convencional
Primavera	4,92	5,10	7,03
Verano	5,25	3,93	5,20
Invierno	4,58	3,71	5,03
Promedio	4,92	4,25	5,75

Tabla 2.3. Precisiones de las Metodologías usadas en el SSEE de Oklahoma

En la estación de primavera los modelos ARIMA son los que producen el menor error, en verano e invierno los modelos ARIMAX son los de menor error, esto se debe a la correlación entre la demanda y la temperatura que existe.

2.3. MÉTODOS DE PRONÓSTICO

Hipótesis de Continuidad

Este supuesto se adopta para los factores externos que influyen las series temporales de la demanda y las variables climáticas, las cuales no deben cambiar para el futuro que se desea pronosticar. Estos factores son:

- Economía en general.
- Competencia en el mercado (oferta).
- Estado del mercado (demanda).
- Estado tecnológico del producto (calidad).

Esta continuidad del ambiente nunca se da en forma perfecta. Se requiere buen juicio para suponer que las violaciones a la continuidad no van a afectar a los resultados de la aplicación del método de pronóstico.

2.3.1. Clasificación de los métodos de pronóstico

En una primera clasificación los métodos de pronóstico se pueden dividir en dos grandes grupos: métodos cuantitativos y métodos cualitativos [32].

a) Métodos Cuantitativos

Pueden ser aplicados bajo tres condiciones:

- 1) Existe información histórica disponible.
- 2) Esta información se puede cuantificar en forma de datos.
- 3) Se puede asumir que el comportamiento histórico se repetirá en el futuro. Esta premisa sirve de base para muchos de los métodos, sean cualitativos o cuantitativos y se aplica especialmente en aquellos métodos que utilizan el comportamiento histórico de la variable a pronosticar como una entrada.

Estos métodos se dividen en dos grupos: métodos intuitivos y métodos formales (basados en técnicas estadísticas).

Métodos Intuitivos: se basan en la experiencia empírica y varían dependiendo de la aplicación, son simples y fáciles de usar pero no siempre muy precisos como los métodos formales.

Métodos Formales o Estadísticos: estos modelos se dividen a su vez en series de tiempo y métodos de regresión.

b) Métodos Cualitativos

Las entradas requeridas dependen del método específico y son principalmente producto del conocimiento, de la capacidad de juicio y de la experiencia adquirida del planificador, por lo que es difícil cuantificar la precisión del pronóstico; su aplicación es casi exclusivamente para mediano y largo plazo. Estos modelos se dividen en dos grupos: métodos explorativos o investigativos y los métodos normativos.

Métodos Explorativos: empiezan con el pasado y presente como punto de inicio y se dirigen hacia el futuro de una manera heurística, analizando todas las posibilidades disponibles.

Métodos Normativos: inician determinando las metas y objetivos futuros y trabajan hacia atrás, analizando si se puede alcanzar dichas metas sujetas a restricciones, recursos y tecnología disponible.

En la bibliografía especializada a los Métodos de PDCP en SSEE se los dividen en dos grupos: los Métodos Estadísticos y las Técnicas basadas en Inteligencia Artificial [31].

2.3.2. Métodos Estadísticos

Están basados en el análisis de series de tiempo y asumen que la demanda es una combinación lineal de variables independientes (modelación paramétrica).

a) Regresión lineal

$$y(t) = a_0 + a_1x_1(t) + \dots + a_nx_n(t) + a(t) \quad (2.17)$$

donde; $y(t)$: demanda
 $x_i(t)$: variables correlacionadas (temperatura, tipo de día, etc.)

$a(t)$: variable aleatoria (media cero y varianza constante)
 a_0, a_1, \dots, a_n : coeficientes de regresión

b) Suavizamiento exponencial

$$y(t) = b^T(t) * f(t) + e(t)$$

$$f(t) = L * f(t - 1)$$
(2.18)

donde; $f(t)$: vector del proceso
 L : Matriz de transición
 $b(t)$: vector de coeficientes
 $e(t)$: ruido blanco

c) Filtro de Kalman y Estimación de Estado

$$x(k + 1) = \phi(k) * x(k) + w(k)$$

$$z(k) = H(k) * x(k) + v(k)$$
(2.19)

donde; $x(k)$: variación del vector de estado
 $w(k)$: ruido blanco con matriz de varianza conocida
 $z(k)$: variación de la medición del vector
 $\phi(k)$: Matriz de transición
 $H(k)$: Matriz de correlaciones
 $v(k)$: error

d) Modelos Econométricos de Box – Jenkins (Tabla 2.4) [5].

AR (auto regresivo)	$y(t) = \phi_1 * y(t - 1) + \dots + \phi_p * y(t - p) + a(t)$ $\phi(B) = 1 - \phi_1 * B - \phi_2 * B^2 - \dots - \phi_p * B^p$
MA (media móvil)	$y(t) = a(t) - \theta_1 * a(t - 1) - \dots - \theta_q * a(t - q) = \theta(B) * a(t)$ $\theta(B) = 1 - \theta_1 * B - \theta_2 * B^2 - \dots - \theta_p * B^p$
ARMA	$y(t) = \phi_1 * y(t - 1) + \dots + \phi_p * y(t - p) + a(t) - \theta_1 * a(t - 1) - \dots - \theta_q * a(t - q)$ $\phi(B) * y(t) = \theta(B) * a(t)$
ARIMA (ARMA integrativo)	$\phi(B) * \nabla^d y(t) = \theta(B) * a(t)$ $\nabla^d y(t) = (1 - B)^d * y(t)$
SARIMA (ARIMA estacional)	$\phi_p(B) * \phi_p(B^S) * \nabla^d \nabla^D y(t) = \theta_q(B) * \theta_q(B^S) * a(t)$
SARIMAX (SARIMA con variables exógenas)	$A(q) * y(t) = B(q) * u(t) + C(q) * a(t)$

Tabla 2.4. Modelos de Box – Jenkins o ARIMA

donde; $y(t)$: demanda al tiempo t
 $u(k)$: temperatura al tiempo t
 B : operador de retardo $\rightarrow y(t - m) = B^m * y(t)$
 p : orden del proceso

$a(t)$: error del pronóstico (ruido blanco)

q : orden del proceso

d : orden del proceso

S : periodo de estacionalidad

2.3.3. Métodos basados en Inteligencia Artificial

Estos modelos tienen la habilidad de mejorar la precisión ya que consideran que la demanda es una función no lineal de las variables intervinientes (demanda, temperatura, etc.). La ventaja es que no requieren ninguna formulación matemática compleja o correlación cuantitativa.

- a) **Sistemas Expertos:** capturan el conocimiento especializado dentro de una base de datos comprensiva, la cual es usada para predecir usando un complejo razonamiento inferencial, explotando el conocimiento de los especialistas para el desarrollo de reglas de inferencia.
- b) **Lógica Fuzzy:** acota un grupo de variables de entrada en una salida usando instrucciones lógicas (if – then). Las reglas fuzzy son obtenidas de los datos históricos usando un algoritmo de aprendizaje.
- c) **Razonamiento Inductivo Fuzzy “Metodología FIR”:** es una mezcla entre la modelación cualitativa (modelación inductiva) y la predicción cuantitativa, esto hace un modelo no paramétrico que no presupone una estructura predefinida.

Se tienen dos tareas principales: la **Modelación Cualitativa** que determina el grupo de variables de entrada que mejor explican el comportamiento entrada-salida de los datos de entrenamiento mediante la identificación de máscaras de correlación (máscaras óptimas fuzzy) y la **Simulación Cualitativa** que compara los datos de prueba con sus k más cercanos vecinos (k -NN “ k Nearest Neighbors”) de la base de datos de entrenamiento e interpola entre las salidas previamente observadas.

Por lo tanto, aprende y analiza los patrones observados de las señales medidas y predice el comportamiento futuro en base de su propio pasado sin identificar el sistema del cual las señales fueron generadas.

- d) **Algoritmos Genéticos.-** representan una poderosa y robusta aproximación para desarrollo heurístico de problemas de optimización combinatorial de gran escala.
- e) **Redes Neuronales Artificiales (ANN).-** son modelos no lineales eficientes, que tienen la capacidad de aprender, identificar y aproximar. Entre las principales tenemos [20]:
 - **Perceptron multicapa:** a estas redes se las conoce también como redes estáticas multicapas alimentadas hacia delante (feedforward).
 - **Red de elementos lineales adaptivos (MADALINE):** constituida por múltiples neuronas ADALINE, cuya estructura tiene un filtro adaptivo en la entrada y utiliza el algoritmo de aprendizaje de Widrow-Hoff (paso descendente) para entrenar su red.

- **Redes Feedforward con retardos temporales (TLFN):**

Enfocada: la memoria de corto plazo es localizada en la parte delantera de la red estática por lo que se limita su aplicación a entornos estacionarios.

Distribuida: usa filtros de respuesta al impulso de duración finita como peso o sinápsis. Tiene la habilidad de hacer frente a la no estacionariedad de la serie mediante una representación temporal distribuida.

- **Redes Recurrentes:**

Red no lineal auto regresiva con variables exógenas (NARX): tiene realimentación de la capa de salida hacia la capa de entrada.

Red recurrente totalmente conectada: tiene realimentación de la capa oculta hacia la capa de entrada.

Perceptron multicapa recurrente con múltiples capas ocultas: tiene realimentación de la salida de cada capa oculta hacia su propia entrada.

Perceptron multicapa recurrente tipo Jordan: tiene realimentación de la capa de salida hacia la capa de entrada.

Red recurrente de segundo orden: usa neuronas de segundo orden.

Se tienen 3 algoritmos de aprendizaje para el entrenamiento de estas redes: backpropagation a través del tiempo (BPTT), aprendizaje recurrente en tiempo real (RTRL) y el Filtro de Kalman desacoplado extendido (DEKF).

Todos los métodos presentados en la bibliografía especializada pueden asociarse, formando modelos híbridos: **Redes Neuronales-Fuzzy, FARMAX (Fuzzy ARMAX), Redes Neuronales-Sistemas Expertos, Algoritmos Genéticos-Redes Neuronales, Redes Neuronales-Filtros de Kalman, Método Bayesiano (probabilidad de Bayes, Redes Neuronales y Regresión), etc.**

En el Estado del Arte se utilizan los modelos econométricos y los modelos ANN como métodos de referencia para validar cualquier nuevo modelo de pronóstico. Por lo tanto en este trabajo se utilizarán estos dos modelos para analizar la performance de la metodología FIR.

2.3.4. Características de los métodos de referencia y FIR

Las características más importantes se muestran en la Tabla 2.5.

ARIMA	ANN	FIR
Modelo lineal.	Modelo no lineal.	Modelo no lineal.
Basados en hipótesis de estacionariedad estricta.	Trabajan mejor el comportamiento no estacionario.	Tratan series no estacionarias como funciones crecientes [36].
Exploran las propiedades del ruido.	Trata de librarse del ruido.	Explora las propiedades estadísticas del ruido.
Problemas para tratar con sistemas multivariable por que incrementan la complejidad de su modelación.	Trabajan bien con sistemas multivariable.	Trabajan mejor con sistemas multivariable.
Requieren mucha intervención y las decisiones que a menudo son tomadas no son triviales (ej. orden del proceso autoregresivo).	Menos sensible a la intervención del usuario pero requiere la selección (heurística) de los parámetros de la red neuronal.	Habilidad para generar modelos aceptables de una manera bastante automática en donde los parámetros necesarios son intuitivos y/o sensibles (ej. número de clases de la demanda).

Tabla 2.5. Características de los métodos de pronóstico

2.4. EL SISTEMA EJEMPLO

El SSEE ecuatoriano que se muestra en la Figura 2.4 será utilizado como Sistema ejemplo para los desarrollos de PDCP del presente trabajo.

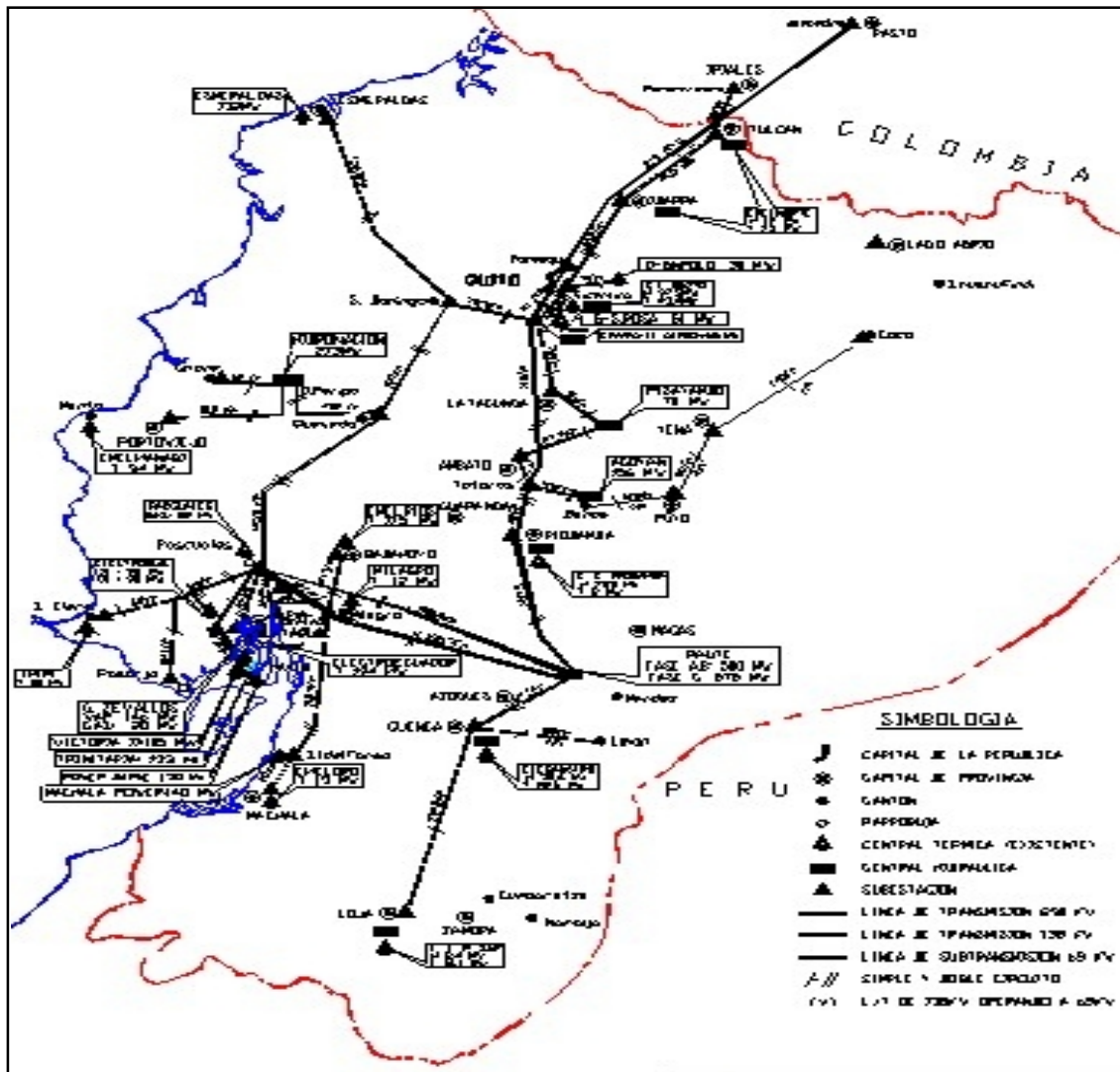


Figura 2.4. SSEE del Ecuador

2.4.1. Características del Sistema ejemplo

El sistema ecuatoriano se divide, a los efectos del PDCP, en dos zonas climáticas: sierra y costa, cada zona con su respectiva temperatura de referencia (Quito y Guayaquil) y tiene las siguientes características:

- Valores horarios de demanda.
- Los precios de la energía se calculan en base a costos variables de producción.
- Demanda máxima: 2.059,8 MW (jueves, 03 de julio de 2003).
- Demanda mínima: 959,7 MW (domingo, 06 de julio de 2003).
- Energía: 35.374 MWh (jueves, 03 de julio de 2003).
- 12 Empresas de Generación.
- 18 Empresas de Distribución (11 disponen de generación y no están escindidas).
- 1 Empresa de Transmisión.
- 22 Grandes Consumidores (demanda mayor a 860 kW).

- Interconexión Colombia – Ecuador en 230 kV y 138 kV.
- El CENACE es responsable de la coordinación técnica y la administración del Mercado Eléctrico Mayorista ecuatoriano, resguardando la seguridad de la operación del Sistema Nacional Interconectado (SNI).

Se utilizará una base de datos que contiene las series históricas de la demanda y la temperatura (sierra y costa) desde el año 2000 hasta el año 2006. En la Figura 2.5 se muestran las series de demanda y temperatura del periodo correspondiente al sábado 09 de enero hasta el viernes 05 de febrero de 2000.

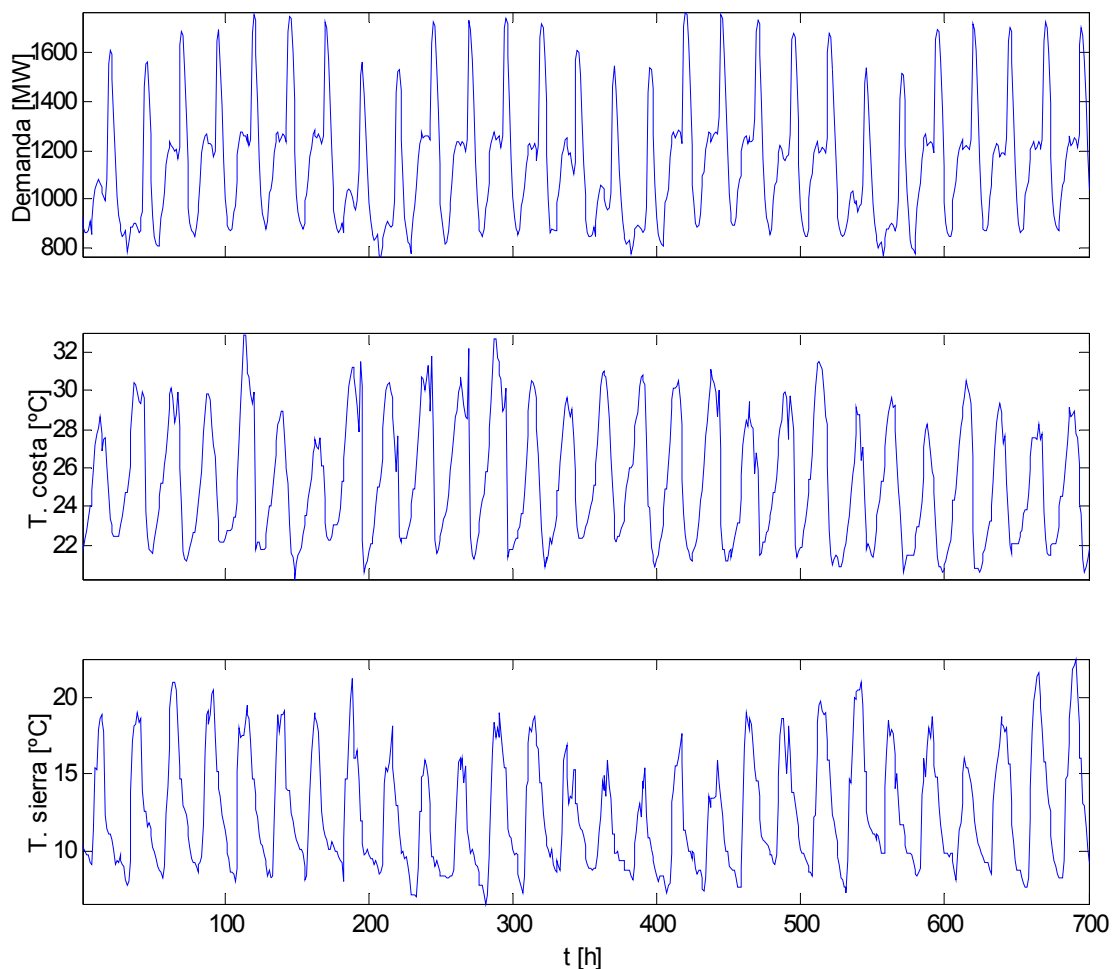


Figura 2.5. Demanda total y temperaturas del SSEE ecuatoriano

2.4.2. Series temporales: demanda y temperatura

La demanda posee estacionalidades: diaria, semanal y anual. Se muestra en la Figura 2.6 la función de autocorrelación de la demanda (Figura 2.5), donde se puede visualizar las estacionalidades: diaria (24 periodos) y semanal (168 periodos).

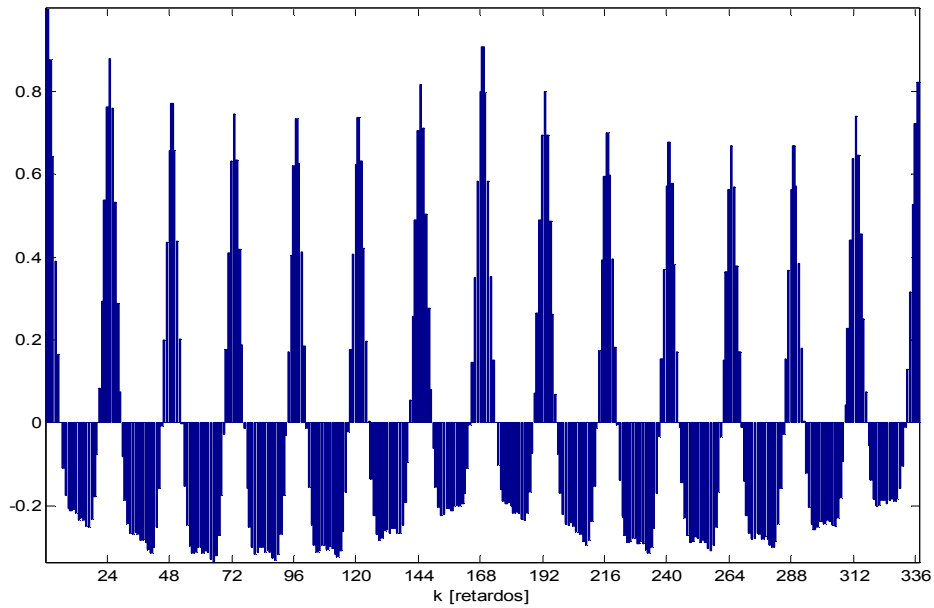


Figura 2.6. Función de Autocorrelación de la demanda

La demanda eléctrica tiene las siguientes características:

- Natural: es un fenómeno físico.
- Estacionaria: en la Figura 2.7 se muestra la media y la desviación estándar de la serie temporal de los días lunes, se analizan los 20 primeros lunes del año 2000. Se presentan 24 gráficas, una para cada hora del día, con lo que se concluye que la media es constante y que la desviación estándar es acotada.

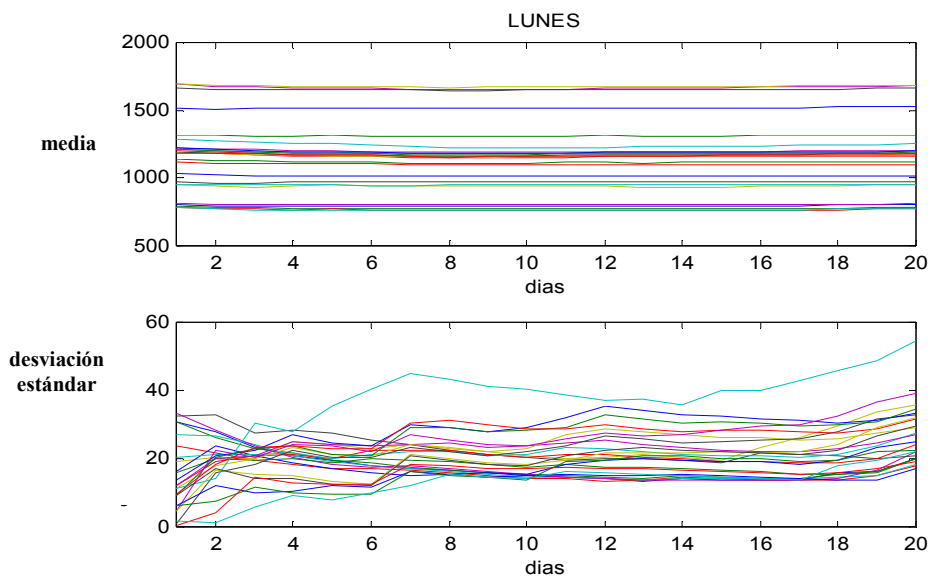


Figura 2.7. Media y la desviación estándar de la serie de los lunes

En la Figura 2.8 se muestra un histograma de la serie correspondiente a los viernes a las 09:00 pm (se analiza los 52 viernes del año 2000).

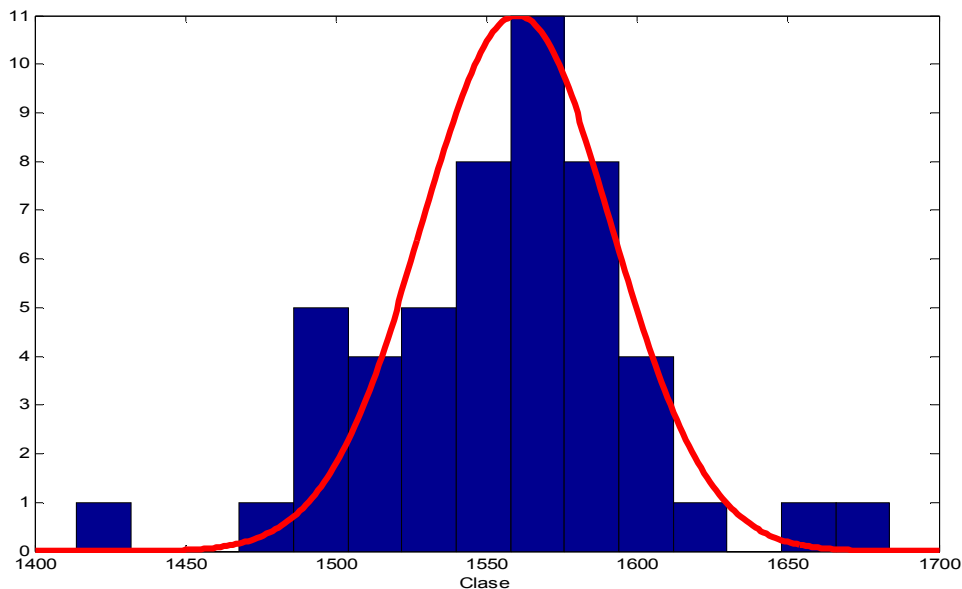


Figura 2.8. Histograma de la serie de los días viernes a las 09:00 pm

Con este análisis se podría considerar que la demanda de los viernes a las 09:00 pm puede ser representada aproximadamente por una función de probabilidad normal.

Método Formal (Test de Fisher) [4]: la detección de la normalizada se realiza estimando la simetría g de la muestra X , y se trata de comprobar la hipótesis $g \approx N(0; 6/N)$. Para un nivel de confianza de 95%, $g < 1,96 * \sigma$, donde $\sigma = \sqrt{\frac{6}{N}}$ y N es la longitud de la serie.

Utilizando este test se demuestra que la serie temporal de la demanda (para cada hora) tiene una distribución normal. Con lo que se concluye que la serie temporal de la demanda es un proceso estacionario.

- Estocástica: intervienen 168 variables aleatorias correspondientes a cada hora de la semana.
- Limpia: se asume que los errores sistemáticos del sistema son despreciables (equipos de medición en buen estado de operación y mantenimiento) y se desprecia el ruido del sistema de medición (sistemas de comunicación confiables). Entonces el ruido en la serie proviene de fuentes de perturbación exteriores.
- No lineal: la demanda es una función no lineal de las variables intervinientes: temperatura, velocidad del viento, humedad, serie histórica de demanda, etc.

- Discreta: serie horaria.

La temperatura se considera una serie temporal sólo con estacionalidad diaria, no se considera la periodicidad climática que se da con periodos anuales, y tiene las siguientes características:

- Natural: es un fenómeno físico.
- Estacionaria: cumple también el test de Fisher.
- Estocástica: muestra un comportamiento aleatorio de los patrones.
- Limpia: se trata a la serie de igual manera que la serie de demanda.
- No lineal: función de las variables exógenas (humedad, zona climática, etc.).
- Discreta: serie horaria.

2.4.3. Modelación de las Incertidumbres en el Sistema ejemplo

- a) **Incertidumbres en la medición de la demanda:** en el SNI la clase de precisión de los transformadores de tensión, de corriente y el medidor de energía normativamente debe ser menor o igual a 0.2% respecto del fondo de escala, es decir: $s_j / j_{max} < 0.002$. Por lo que la incertidumbre asociada a la medición de la demanda en un punto de medición según la ecuación (2.14) debe ser menor o igual a 0.4828%.

La demanda horaria del sistema ecuatoriano es producto de la suma de 35 puntos de medición (puntos de entrega del sistema de potencia a las empresas distribuidoras), por lo que la incertidumbre total de la medición de la demanda será la suma de todas las incertidumbres de los puntos de medición, y por lo tanto menor o igual a 16.9%, ya que se considera que las observaciones son independientes (no correlacionadas).

La medición de la temperatura se ve afectada solo por los errores de la termocupla y la precisión de estos instrumentos de medición es de 0.2 % (incertidumbre).

Se desprecian los errores de transmisión de datos (ruido del sistema de medición).

- b) **Incertidumbres en la obtención de datos:** la Dirección de Operaciones del CENACE es la responsable del análisis pos-operativo del sistema, que utiliza los valores promedio de los días similares de la serie histórica para corregir los datos atípicos ocasionados por fallas en el sistema de potencia, errores humanos en la recolección de datos, etc.
- c) **Incertidumbres de los factores que influyen la demanda:** el PDCP realizado por el CENACE incorpora solo datos de la serie temporal de la demanda, por lo que se supone existen incertidumbres debido a la no modelación de la temperatura como variable externa, ya que se ha demostrado que el mayor centro de consumo de la costa (Guayaquil) es una zona en la que existe correlación entre la demanda y la temperatura de 0.6 (ver Figura 2.9) [43].

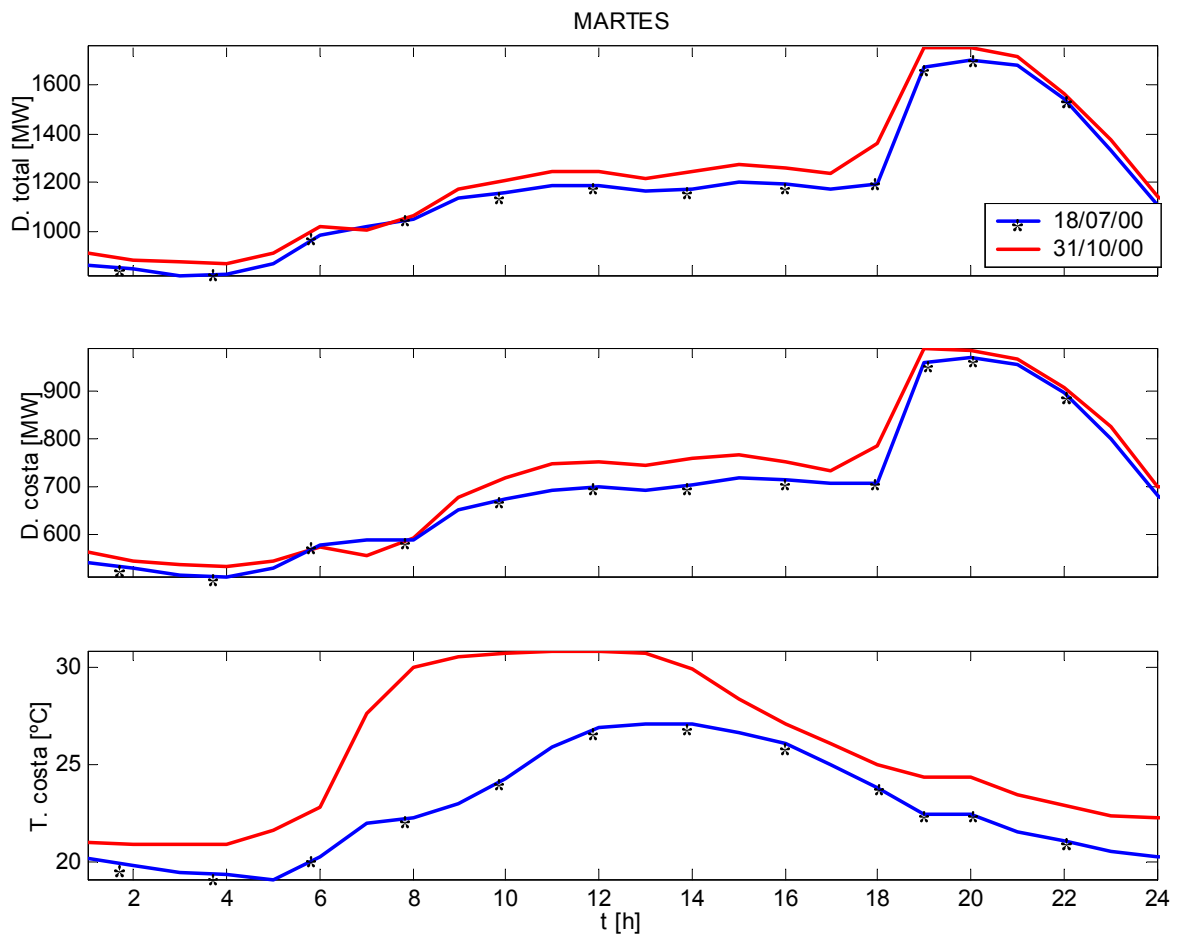


Figura 2.9. Influencia de la temperatura en el Sistema ejemplo

2.4.4. Pronóstico de demanda utilizando modelos econométricos

La Dirección de Planeamiento del CENACE diariamente es responsable de obtener el PDCP para un día y una semana de adelanto, como resultado de la aplicación de un Modelo SARIMA dado por la ecuación (2.20).

$$(1 - B^{24}) \prod_p (1 - \phi_i B^i) y_{(t)} = \prod_q (1 - \theta_q B^q) e_{(t)} \tag{2.20}$$

- donde;
- $y_{(t)}$: demanda horaria al tiempo t
 - B : operador de retardo
 - ϕ_p : parámetros del proceso auto regresivo (AR)
 - θ_q : parámetros del proceso de media móvil (MA)
 - $e_{(t)}$: error del pronóstico (ruido blanco)

Se desarrolla la ecuación (2.20) para determinar cuáles son las entradas del modelo, considerando que se modela procesos AR de orden 1 y 24 y un proceso MA de orden 24 (ecuación (2.21)).

$$y_{(t)} = y_{(t-24)} + \phi_1 [y_{(t-1)} - y_{(t-25)}] + \phi_{24} [y_{(t-24)} - y_{(t-48)}] - \phi_1 \phi_{24} [y_{(t-25)} - y_{(t-49)}] + (1 - \theta_{24} B^{24}) e(t) \quad (2.21)$$

El esquema del modelo de pronóstico (de acuerdo con la ecuación (2.21)), se muestra en la Figura 2.10.

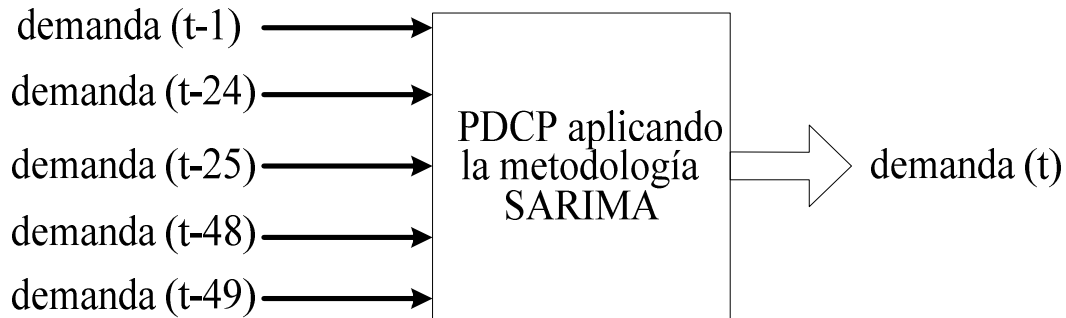


Figura 2.10. Esquema del PDCP utilizando modelos SARIMA

2.4.5. Pronóstico de demanda utilizando redes recurrentes de Jordan

En la referencia [21] se utilizan las redes recurrentes de Jordan para el PDCP del sistema ejemplo.

En la referencia se analiza el periodo desde el 01 de junio hasta el 31 de octubre de 2000. En el periodo de junio a septiembre se seleccionan aleatoriamente las duplas que forman cada uno de los conjuntos de entrenamiento, validación y test. El modelo óptimo encontrado tiene la estructura [15 60 30 1], es decir 15 neuronas en la capa de entrada, 60 neuronas en la primera capa oculta, 30 neuronas en la segunda capa oculta y 1 neurona en la capa de salida. El grafo se muestra en la Figura 2.11.

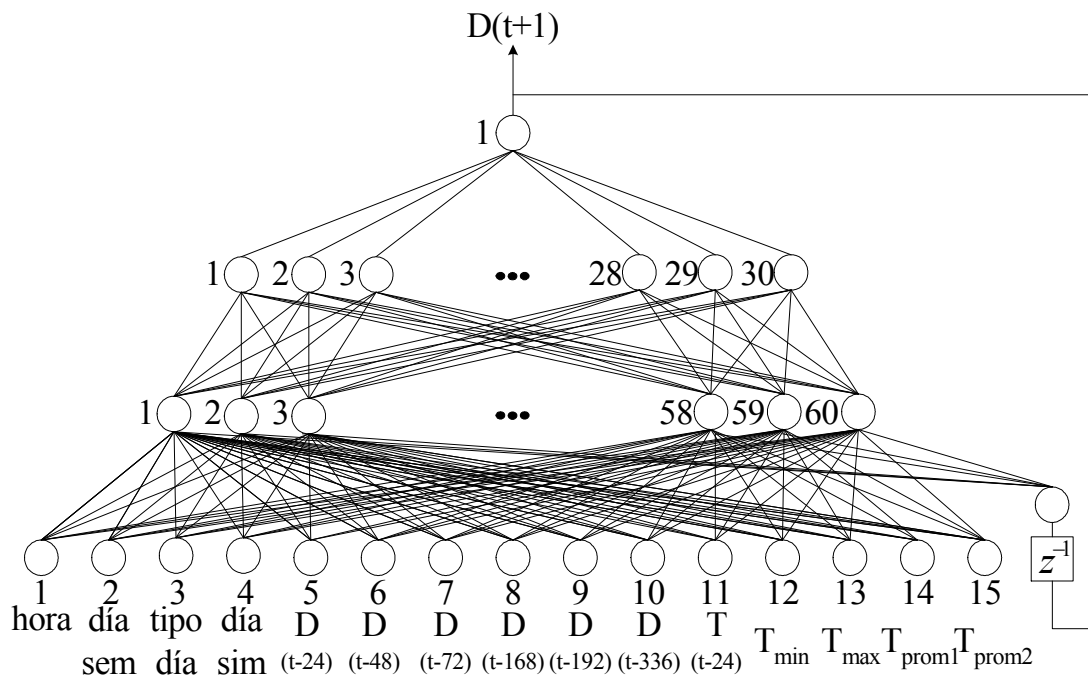


Figura 2.11. Red recurrente de Jordan utilizada para el PDCP

Arquitectura de la red

- La función de activación de las capas ocultas es del tipo tangente hiperbólico y la función de activación de la capa de salida es lineal.
- Las variables explicativas y la variable explicada son escaladas en el intervalo $[-1,1]$.
- La inicialización de los pesos y bias de cada capa se realiza aleatoriamente en el intervalo $[-1,1]$ con el algoritmo de Nguyen-Widrow [41].

Este algoritmo elige los valores a fin de distribuir las regiones activas de cada neurona equitativamente en el espacio de entrada, disminuyendo el tiempo de entrenamiento.

- El algoritmo de aprendizaje utilizado es el Gradiente Conjugado Escalado [35] desarrollado por Moller, el cual asegura mayor rapidez en el proceso de convergencia comparado con otros algoritmos de entrenamiento. Se utiliza el error cuadrático medio (mse) para analizar la performance del algoritmo.

Armado de duplas de Entrada – Salida

Se arman las duplas (patrones de entrada – salida) constituidas por los datos de las variables explicativas de tal manera que el conjunto de entrenamiento contenga por lo menos 4/5 de las duplas representativas y el conjunto de test se aconseja contenga el 1/5 del total de duplas para asegurar la capacidad predictiva de los ejemplos abarcados por los datos observados.

Las redes neuronales artificiales tienen problemas de sobreentrenamiento en la etapa de aprendizaje y esto sucede cuando el método ha memorizado las muestras de entrenamiento y no puede generalizar con nuevas muestras (datos de test). El método de solución “**Early Stopping**” que se propone, consiste en separar las series temporales en 3 grupos de datos:

- **Datos de entrenamiento:** es el grupo de datos que se utiliza para determinar el modelo de pronóstico, los parámetros de la estructura, calcular el gradiente descendente, actualizar los pesos de la red neuronal, encontrar la máscara de correlación, etc.
- **Datos de validación:** El error de estos datos se monitorea durante el proceso de entrenamiento y cuando el método de pronóstico utilizado empieza a memorizar los datos de entrenamiento el error de validación empieza a aumentar, por lo que el entrenamiento debe parar cuando ocurra esta situación.
- **Datos de test:** es el grupo de datos que se quiere pronosticar y que servirán para determinar el error del pronóstico. Estos datos no se usan en el entrenamiento pero sirven para comparar diferentes modelos.

La selección de los grupos de entrenamiento y validación es de tal manera que el conjunto de entrenamiento y el de validación contengan cada uno al menos 2/5 del total de duplas representativas.

La red se entrena utilizando 600 ciclos, tomando como criterio de parada el método Early Stopping. En la Figura 2.12 se muestra la curva de entrenamiento y validación de la red óptima.

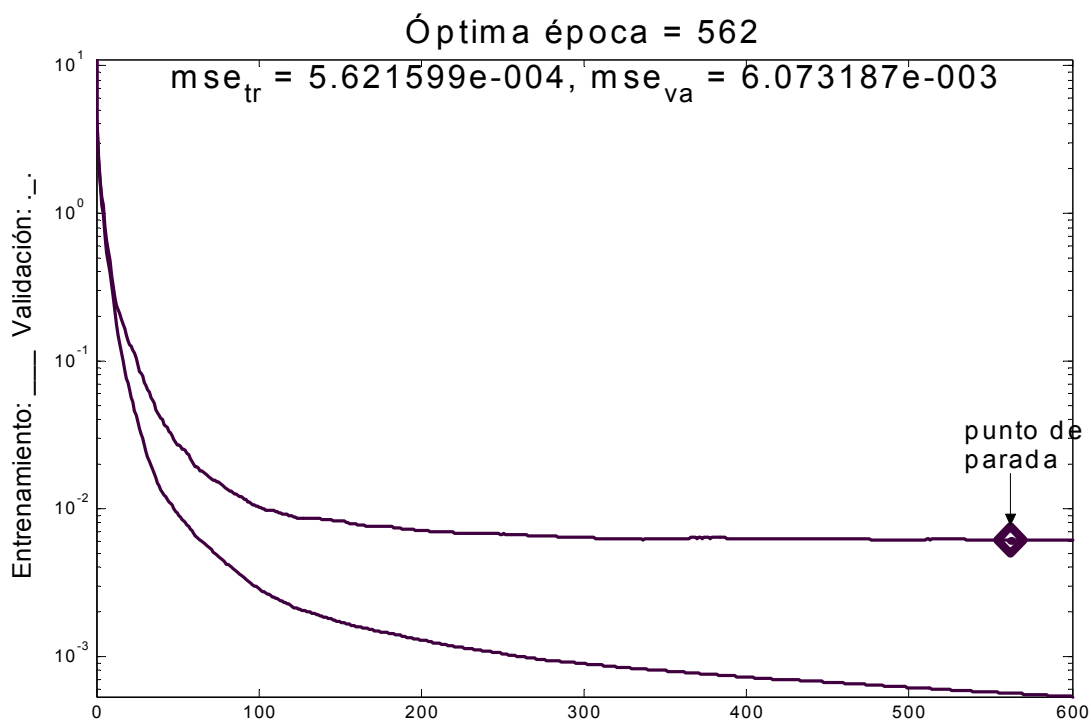


Figura 2.12. Aprendizaje de la red recurrente

En la Figura 2.13 se muestra la demanda real y el PDCP para el mes de octubre, se incluye la ACF de los errores de pronóstico, ya que el **Test de independencia de Anderson** [4] acepta que el proceso es puramente aleatorio (ruido blanco) cuando la ACF se distribuye según una función normal $N(0,1/n)$, donde n es el tamaño de la muestra.

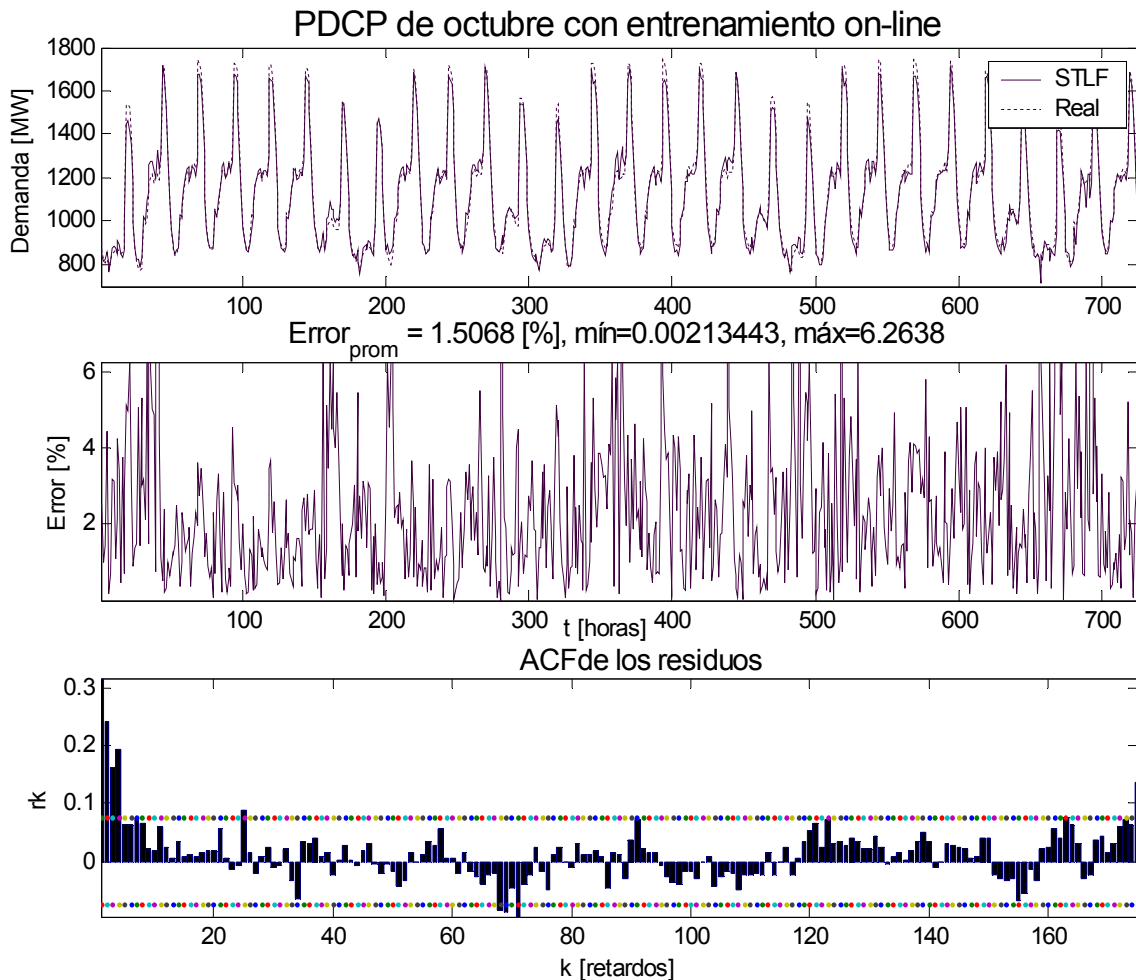


Figura 2.13. PDCP del sistema ejemplo para el mes de octubre

Se observa que para un límite de significación del 95% el error de pronóstico aproxima a un proceso aleatorio.

En la Tabla 2.6 se resume los errores relativos promedios del PDCP, incluyéndose también los errores de pronóstico con los Modelos SARIMA (utilizados actualmente en el CENACE), las redes feedforward (MLP) y las redes neuronales recurrentes (RR).

PDCP de octubre	Error [%]
SARIMA*	1,6640
MLP** [15 40 20 1]	1,6911
MLP [15 60 30 1]	1,6312
MLP [15 80 40 1]	1,7492
RR [15 60 30 1]	1,5068

Tabla 2.6. Errores del PDCP del sistema ejemplo

* Dato del sistema de prueba.

** Redes Feedforward, normalmente utilizadas en el Estado del Arte.

3. METODOLOGÍA DE SOLUCIÓN PROPUESTA: RAZONAMIENTO INDUCTIVO FUZZY

El razonamiento humano ha sido entendido como un proceso mental de simulación debido a que los humanos somos capaces de tomar decisiones aún sin tener el conocimiento total del fenómeno solo reconociendo los patrones pudiendo así analizar el comportamiento del sistema.

3.1. RAZONAMIENTO INDUCTIVO FUZZY (FIR)

La metodología FIR (Fuzzy Inference Reasoning) es una implementación de la metodología de análisis general de sistemas “General System Problem Solver – GSPS” [27, 28 y 47], que posibilita estudiar los modos conceptuales de comportamiento de sistemas dinámicos.

Fue reimplementada por F. Cellier de la Universidad de Arizona [7]. Mediciones Fuzzy fueron añadidas a la metodología de modelación para tratar las incertidumbres de los límites [29], donde también se incluye el tratamiento de valores perdidos y las mediciones para estimar la predicción del error. Este proceso de razonamiento es realizado usando el programa SAPS-II [11].

Esta metodología previamente ha sido aplicada a la modelación y control de sistemas dinámicos [30, 33, 36, 37, 39 y 48] con buenos resultados. En la Figura 3.1 se muestran las cuatro etapas de la metodología:

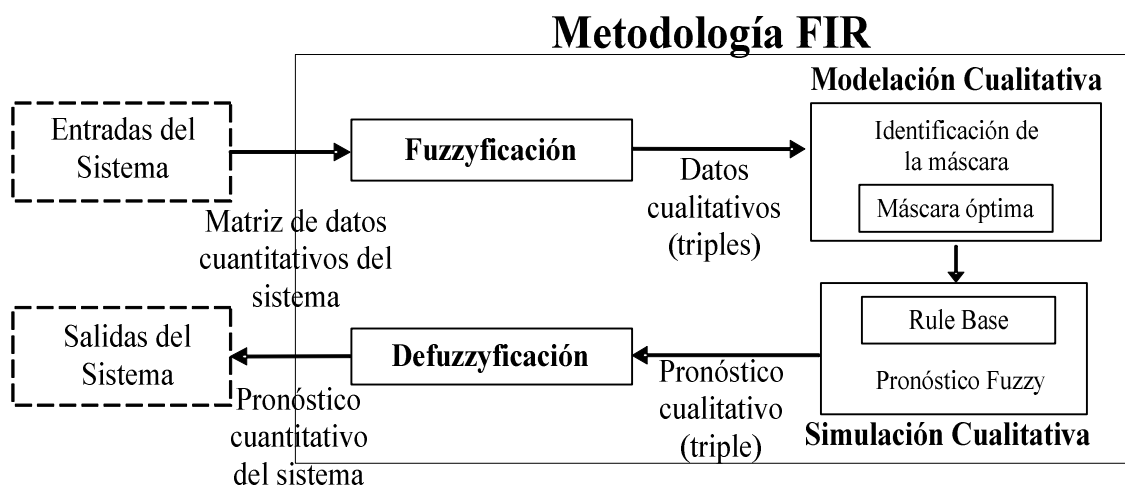


Figura 3.1. Etapas de la Metodología FIR

Es importante notar que el razonamiento propiamente dicho no involucra las etapas de fuzzyficación y defuzzyficación, pero son esenciales para que la metodología FIR opere en un entorno mixto (cualitativo/cuantitativo) de modelación y simulación.

3.1.1. Fuzzyficación (Fuzzy recoding)

El objetivo de este proceso es aumentar drásticamente la velocidad de optimización. Si tenemos una relación entre n entradas y una salida, en lugar de realizar la búsqueda en el espacio continuo n -dimensional para encontrar el óptimo patrón de entrada/salida, la búsqueda es limitada al espacio discreto c -dimensional de los valores clase [9].

Este proceso convierte los datos cuantitativos que describen el comportamiento del sistema (series temporales: demanda y temperatura y variables auxiliares: tipo de día, hora, etc.) a una trayectoria triple cualitativa.

El primer elemento de la variable triple se llama “valor clase” (class value) y representa una discretización gruesa de los valores originales de la variable. El segundo elemento se llama “valor de membresía fuzzy” (fuzzy membership value) y denota el nivel de confianza expresado en el valor clase elegido para representar un particular valor cualitativo. El tercer elemento se llama “valor del flanco” (side value) [8], el cual indica si el valor cuantitativo esta a la derecha o a la izquierda del valor pico de la función de membresía fuzzy asociada. El valor del flanco que es particular de la metodología FIR (no es comúnmente utilizado en la lógica fuzzy), es responsable de preservar el completo conocimiento del valor original cuantitativo

Un ejemplo de la fuzzyficación de la temperatura se ilustra en la Figura 3.2. Donde los valores clase pueden ser reemplazados por números enteros (frío=1, ..., caliente=5).

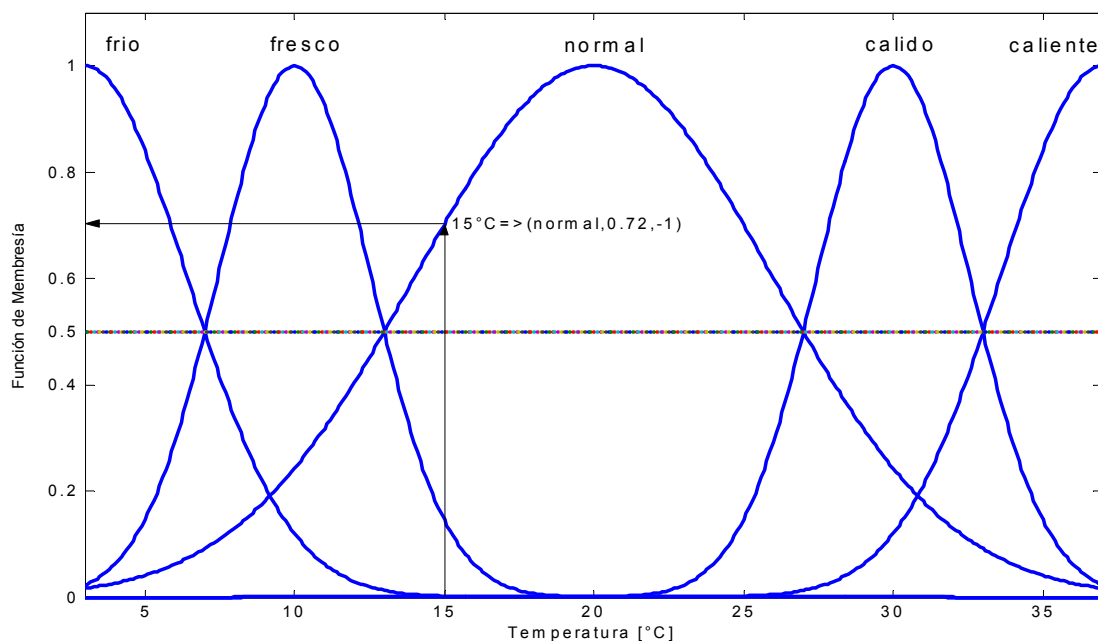


Figura 3.2. Fuzzyficación de la temperatura

Las funciones gaussianas de membresía fuzzy pueden ser calculadas usando la ecuación (3.1).

$$Memb_i = e^{-\tau_i(x-\mu_i)^2} \quad (3.1)$$

donde; x : trayectoria continua a ser fuzzyficada
 i : índice de la curva gaussiana (representa las clases seleccionadas)
 μ_i : media (centro del intervalo de la clase elegida)
 τ_i : es determinado, tal que la función de membresía $Memb_i$ baja a 0,5 en los límites del intervalo asociado con la clase elegida [29]

Contrariamente a otras aproximaciones fuzzy, las colas de las funciones de membresía ($Memb_i < 0,5$) son ignoradas, esto se justifica con la selección de la técnica de inferencia fuzzy [38].

En este punto, la pregunta que se debe responder es cuantas clases deberían ser seleccionadas para cada variable y cuáles serían los límites para cada clase. El número de clases elegido para cada variable influencia directamente la complejidad del cálculo de la etapa de inferencia, por lo que su selección podría ser crítica. Usualmente los sistemas fuzzy tradicionales requieren entre 7 y 13 clases para cada variable.

La selección de la técnica de inferencia hace posible reducir el número de niveles entre 3 y 5, rango confirmado por una serie de aplicaciones prácticas de la metodología FIR [30, 33, 36, 37, 39 y 48].

Procedimiento para la determinación del número de clases y sus límites

El número de clases elegido para cada variable es muy importante, ya que determina la expresividad y predictividad del modelo cualitativo. La **expresividad** de un modelo es una medida de la información contenida que el modelo provee y la **predictividad** es una medida de su capacidad de pronóstico [29].

Si las variables se convierten a una sola clase, el sistema exhibirá solamente un estado factible (Modelo Nulo) y es capaz de predecir infinitamente el comportamiento del sistema. Este modelo se caracteriza por una infinita predictividad y una expresividad nula.

Por otro lado si cada variable se fuzzyfica dentro de por ejemplo 1.000 clases, el sistema tendrá un gran número de estados factibles, por lo que la expresividad del modelo será excelente pero la predictividad será pobre (a menos que se cuente con una gran cantidad de observaciones), resultando la selección del número de clases una solución de compromiso. El procedimiento para la estimación del número de clases es el siguiente:

- 1) Como primera medida se puede consultar a especialistas sobre el número de clases en que se puede modelar cada variable.
- 2) Cuando no se tiene una idea clara del número de clases (clusters) en que deberían ser fuzzyficadas las series, se calcula el histograma de la serie temporal para cada variable del sistema. Este histograma provee información sobre la distribución

“natural” de los datos, por lo que el planificador se ayuda de éste para decidir el número de clases en que las variables del sistema serán discretizadas.

- 3) Se puede aplicar también el algoritmo de Clasificación “Subtractive Clustering”, el cual estima el número de clusters y sus respectivos centroides, con lo que se puede contrastar la información que se deriva del histograma. En el apéndice A se explica detalladamente el algoritmo.

Del histograma se puede hacer una primera estimación visual (gruesa) de los límites de cada clase, pero hay varios métodos formales para encontrar estos límites: “Equal Frequency Partition”, “Equal Width Partition”, “K-Means” y “Fuzzy C-Means” (apéndice B). El método de clustering Fuzzy C-Means [26], se utiliza en esta investigación, por ser el que mejor estima en el histograma los límites de cada clase y se compara con el método utilizado por la metodología FIR tradicional (Equal Frequency Partition) [8]

En el Estado del Arte para la modelación del PDCP se consideran variables exógenas como la temperatura y el tipo de día (normal de trabajo, fin de semana, día feriado, etc.). La fuzzyficación, selección y modelación cualitativa de las mismas se verá en el ítem 3.2.

3.1.2. Modelación cualitativa (Qualitative Modeling)

La modelación cualitativa tiene como objetivo identificar los patrones (espaciales y temporales) en el tiempo, mediante la determinación del grupo de variables de entrada que mejor explican el comportamiento entrada – salida. Tal relación es llamada máscara de correlación (mask).

Búsqueda de la Máscara Óptima

- a) En la metodología FIR definida por Cellier, las matrices candidatas a máscara óptima son un conjunto de máscaras posibles, de las cuales la mejor es elegida por un mecanismo de búsqueda exhaustiva (Problema Combinatorial). La matriz candidata contiene -1 en donde la máscara tiene una potencial entrada, $+1$ donde la máscara tiene su salida y 0 donde se considera que la entrada no tiene ninguna correlación con la salida.

El algoritmo de búsqueda exhaustiva empieza investigando a través de todas las máscaras de complejidad dos (máscaras con una sola entrada y una salida) y encuentra la máscara con el mayor valor de calidad, esta calidad es definida en base a la Entropía de Shannon. Luego procede a buscar todas las máscaras de complejidad tres (máscaras con dos entradas y una salida) y encuentra la de mayor valor de calidad y continúa de la misma manera hasta alcanzar la máxima complejidad (número máximo de las m -entradas + 1).

La complejidad de la máscara primero irá creciendo con el incremento de la complejidad, luego alcanza un máximo y decae rápidamente, heurísticamente un buen valor para la máxima complejidad es usualmente 5 o 6 [10].

La Entropía de Shannon [10], se utiliza para determinar la incertidumbre asociada al patrón de entrada – salida.

- b) En esta investigación se plantea resolver un problema de optimización para encontrar máscaras de correlación que tienen un mejor comportamiento comparadas con las máscaras encontradas por la metodología FIR tradicional, como se muestra en el capítulo 5.

3.1.3. Simulación cualitativa (Qualitative Simulation)

El Objetivo de esta etapa es analizar y aprender patrones observados en la máscara de correlación con el fin de predecir el comportamiento futuro en base a los datos históricos.

Algoritmo: la máscara se ubica al final de la matriz de valores clase y se extraen las m -entradas (patrón de valores clase), los cuales son comparados en la matriz clase de comportamiento para obtener los valores clase futuros de la salida. Además, se extraen los valores de membresía y flanco de los estados correspondientes.

A los efectos de la comparación con los valores históricos de la matriz clase, los estados de entrada obtenidos con la máscara son transformados mediante una pseudo regeneración (normalización), como se detalla en la ecuación (3.3).

$$p_i = Side_i * B * \sqrt{-\ln(Memb_i)} + 0,5 \quad (3.3)$$

donde; $B = (-4 * \ln(0,5))^{-\frac{1}{2}}$

Para las clases extremas los p_i son calculados mediante la ecuación (3.4).

$$\begin{aligned} \text{Para la primera clase:} & \quad p_i = C * \sqrt{-\ln(Memb_i)} \\ \text{Para la última clase:} & \quad p_i = 1 - C * \sqrt{-\ln(Memb_i)} \end{aligned} \quad (3.4)$$

donde; $C = (-\ln(0,5))^{-\frac{1}{2}}$

Los valores posición p_i son variables cuantitativas que pueden ser usados para representar la magnitud relativa de un estado cualitativo (triple). Los valores p_i de los estados de entrada son concatenados en el vector p_{in} (ecuación (3.5)).

$$p_{in} = [p_1, p_2, \dots, p_m] \quad (3.5)$$

De igual manera se calculan los p_i para cada registro obtenido de la matriz de comportamiento cuyas m -entradas son idénticas a las del patrón de entrada, llamando a este vector p_{in}^k , (estos vectores no son iguales ya que poseen funciones de membresía y flanco diferentes).

Finalmente, la norma euclidiana R^2 es calculada mediante la ecuación (3.6).

$$d_k = \|p_{in} - p_{in}^k\|_2 = \sqrt{\sum_{i=1}^N (p_{i,in} - p_{i,in}^k)^2} \quad (3.6)$$

Se identifican los 5 p_{in}^k , con menor distancia d_k (si se encuentran al menos 5). Este grupo de elementos es llamado 5-NN. La contribución de cada vecino a la estimación de la predicción es función de su proximidad y es expresada por un promedio ponderado de los valores de salida, donde los pesos w_{rel} se calculan de acuerdo a la distancia d_k .

Los pesos absolutos w_{abs} son calculados mediante la ecuación (3.7).

$$w_{abs_k} = \left(\frac{d_{max} - d_k}{d_{max}} \right)^x \quad (3.7)$$

donde; $d_{max} = \max\{d_k\}$ para todo k elemento de los 5-NN.
 x : factor que permite ponderar a los más cercanos vecinos.

Los pesos relativos se calculan mediante la ecuación (3.8).

$$w_{rel_k} = \frac{w_{abs_k}}{\sum_{\forall k} w_{abs_k}} \quad (3.8)$$

Los pesos relativos se puede interpretar como porcentajes, por lo que el nuevo estado de salida o puede ser calculado como el promedio ponderado de los estados de salida (o_k) de los 5-NN, Entonces el estado de salida cualitativo triple se calculan por medio de la ecuación (3.9).

$$o = \sum_{\forall k} w_{rel_k} * o_k \quad (3.9)$$

Los valores de membresía y flanco de la salida pueden ser obtenidos mediante la función inversa de la pseudo regeneración planteada en la ecuación (3.3).

3.1.4. Defuzzyficación (Regeneration)

En esta etapa las predicciones de los valores de membresía, funciones flanco y clase de la salida promedio son convertidas a predicciones cuantitativas usando la operación inversa a la fuzzyficación.

3.1.5. Ejemplo de aplicación de la metodología FIR

Se tiene un sistema lineal descrito por las siguientes ecuaciones dinámicas:

$$\begin{bmatrix} \bullet \\ x_1 \\ \bullet \\ x_2 \\ \bullet \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -3 & -4 \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} * u$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} * u$$

Calculando la respuesta del sistema a la función paso unitaria, obtenemos el tiempo de establecimiento $t_l \approx 6s$, se decide usar $\delta t \approx 3s$, con lo que la profundidad de la máscara será 3. Una señal aleatoria binaria es usada como entrada en el sistema, esto es para poder excitar de una manera óptima todas las frecuencias relevantes.

El proceso de fuzzyficación es aplicado a cada trayectoria observada, obteniéndose 3 matrices de igual dimensión, en donde, cada una de las columnas de estas matrices representa una variable observada y cada fila representa un estado registrado, el proceso se muestra en la Figura 3.3.

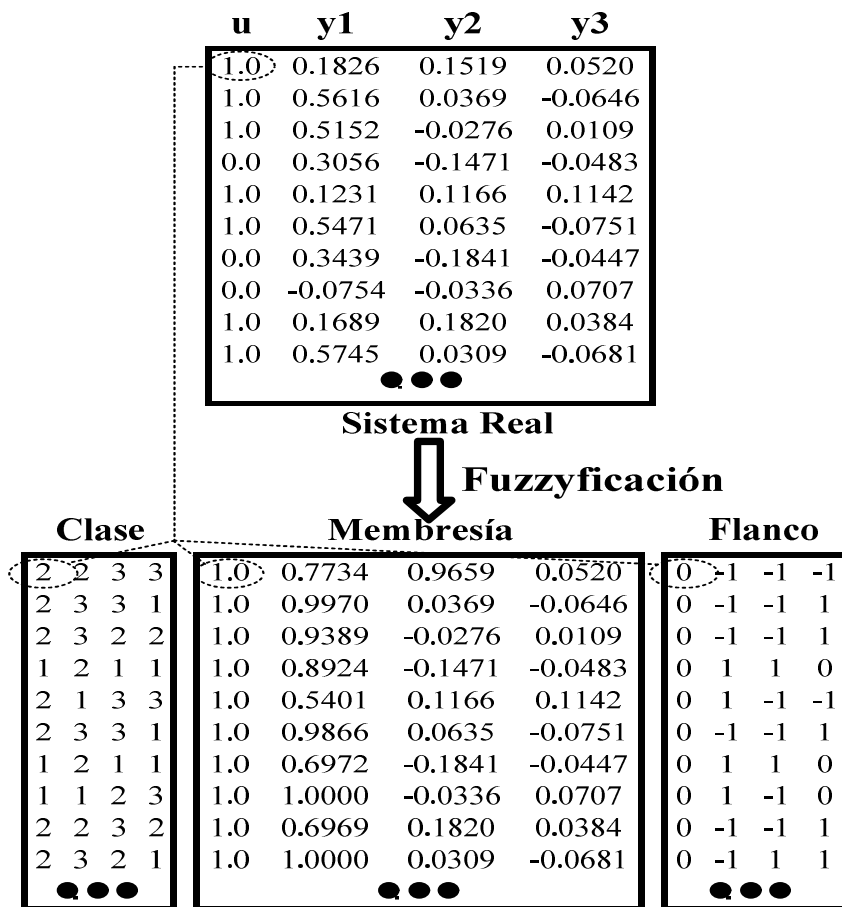


Figura 3.3. Fuzzyficación del sistema lineal

Una posible relación entre las variables cualitativas podría estar dada por la ecuación (3.2).

$$y_1(t) = \tilde{f}(u_1(t-2t), y_3(t-2t), u_1(t-\delta t), u_1(t)) \quad (3.2)$$

donde; \tilde{f} : denota una relación cualitativa genérica.

La ecuación anterior se utiliza para determinar la salida y_1 y puede ser representada por la siguiente matriz (máscara de correlación). Los elementos negativos representan entradas y el valor de +1, representa la salida del modelo.

$$\begin{array}{c|cccc} t \setminus x & u_1 & y_1 & y_2 & y_3 \\ \hline t - 2\delta & -1 & 0 & 0 & -2 \\ t - \delta & -3 & 0 & 0 & 0 \\ t & -4 & 1 & 0 & 0 \end{array}$$

Una vez que la máscara óptima ha sido determinada, esta es aplicada a la matriz de códigos de clase, obteniéndose una particular matriz de valores clase de entrada/salida, esta muestra para cada estado de entrada, cual salida es más probable de ser observada. El mismo proceso se hace con las matrices de membresía y valores flanco, obteniéndose las correspondientes matrices de entrada/salida.

En la Figura 3.4 se muestra el proceso de pronóstico para la salida y_1 del sistema lineal.

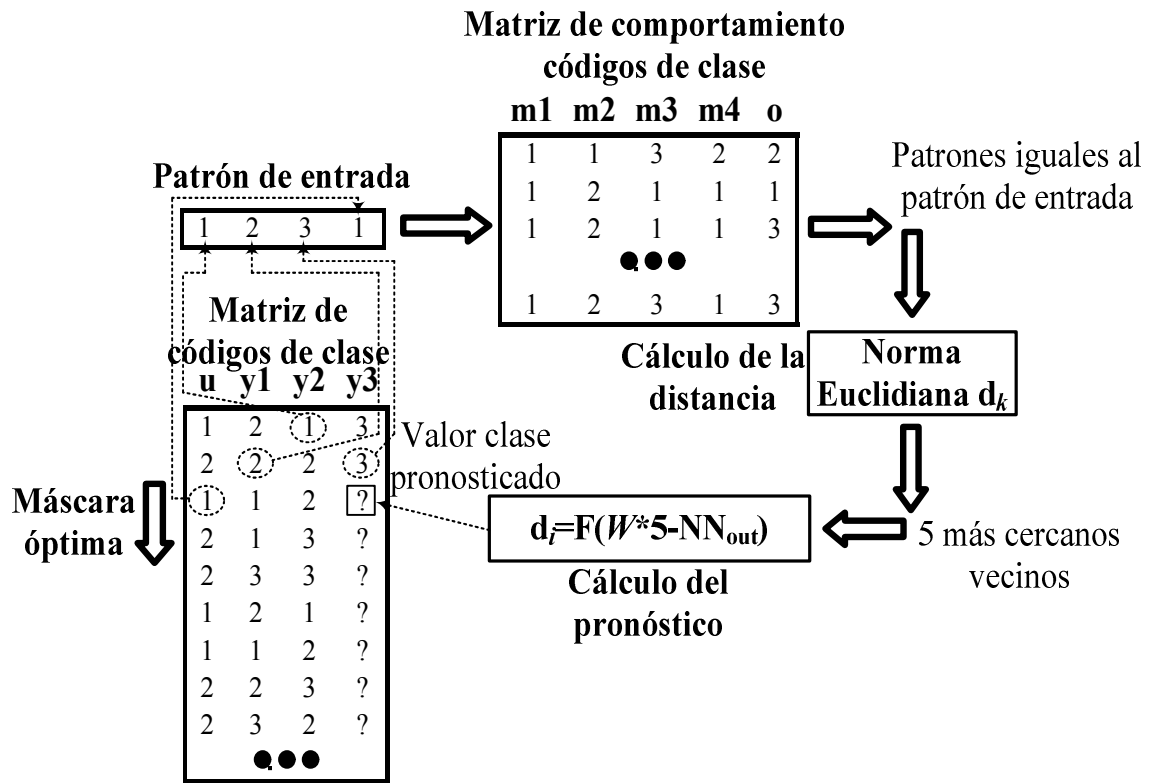


Figura 3.4. Diagrama del proceso de Pronóstico Fuzzy

Cuando aplicamos la máscara a los datos cualitativos, los valores de las m -entradas son extraídos y la matriz de comportamiento es usada para determinar el valor futuro de la salida. Los 5 registros previos con la menor norma R^2 son identificados (si al menos 5 registros son encontrados de la matriz clase de comportamiento) y son usados para pronosticar el nuevo estado de salida.

Los resultados de la simulación utilizando la metodología FIR se muestran en la Figura 3.5. La simulación se realizó utilizando SAPS II.

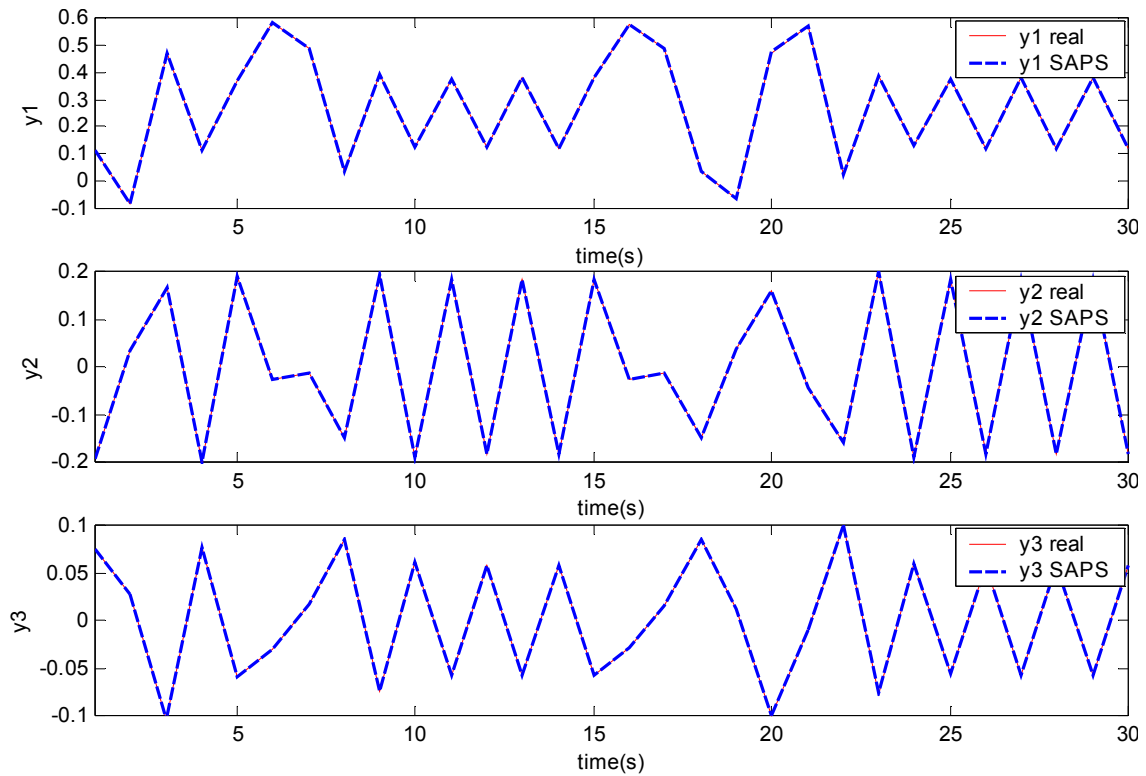


Figura 3.5. Señales reales y pronosticadas del Sistema lineal

El error cuadrático medio en porcentaje es 0.0014% para la señal y_1 , 0.0366% para la señal y_2 y 0.0324% para la señal y_3 .

3.2. APLICACIÓN DE LA METODOLOGÍA FIR AL PRONÓSTICO DE DEMANDA EN EL SISTEMA ECUATORIANO

En el Estado del Arte para la modelación del PDCP se incluyen variables auxiliares como temperatura, día normal de trabajo, fin de semana, día feriado, etc. En la Tabla 3.1 se muestra la estructura del sistema multivariable del problema a plantear para el PDCP en esta investigación, donde se observan las series temporales que se consideran.

Tipo de día [A]	T_min sierra [B]	T_max sierra [C]	T_min costa [D]	T_max costa [E]	Demanda [F]
day (t-n)	Tm_Uio (t-n)	TM_Uio (t-n)	Tm_Gye (t-n)	TM_Gye (t-n)	load (t-n)
day (t-n-1)	Tm_Uio (t-n-1)	TM_Uio (t-n-1)	Tm_Gye (t-n-1)	TM_Gye (t-n-1)	load (t-n-1)
...
day (t-1)	Tm_Uio (t-1)	TM_Uio (t-1)	Tm_Gye (t-1)	TM_Gye (t-1)	load (t-1)
day (t)	Tm_Uio (t)	TM_Uio (t)	Tm_Gye (t)	TM_Gye (t)	load (t)

Tabla 3.1. Series temporales para el PDCP

3.2.1. Fuzzyficación

Se procede a realizar el proceso de fuzzyficación de cada una de las variables, utilizando los criterios anteriormente citados (ítem 3.1.1).

a) Tipo de día

Esta variable se la convierte en cuatro clases (lunes, martes a viernes, sábado y domingo). El proceso de transformación se muestra en la Tabla 3.2.

día	clase	codificación	membresía	flanco
lunes	Laboral 1	1	1	0
martes	Laboral 2	2	1	0
miércoles	Laboral 2	2	1	0
jueves	Laboral 2	2	1	0
viernes	Laboral 2	2	1	0
sábado	Semilaboral	3	1	0
domingo	No laboral	4	1	0

Tabla 3.2. Fuzzyficación de la variable Tipo de día

El día lunes (laboral 1) se lo diferencia de los demás días normales de trabajo (laboral 2), ya que este día posee un factor inercial en las horas de la mañana debido a la terminación del fin de semana (demanda menor a la del resto de días).

Las funciones que se utilizan para esta transformación (funciones trapezoidales) de la variable tipo de día se muestran en la Figura 3.6.

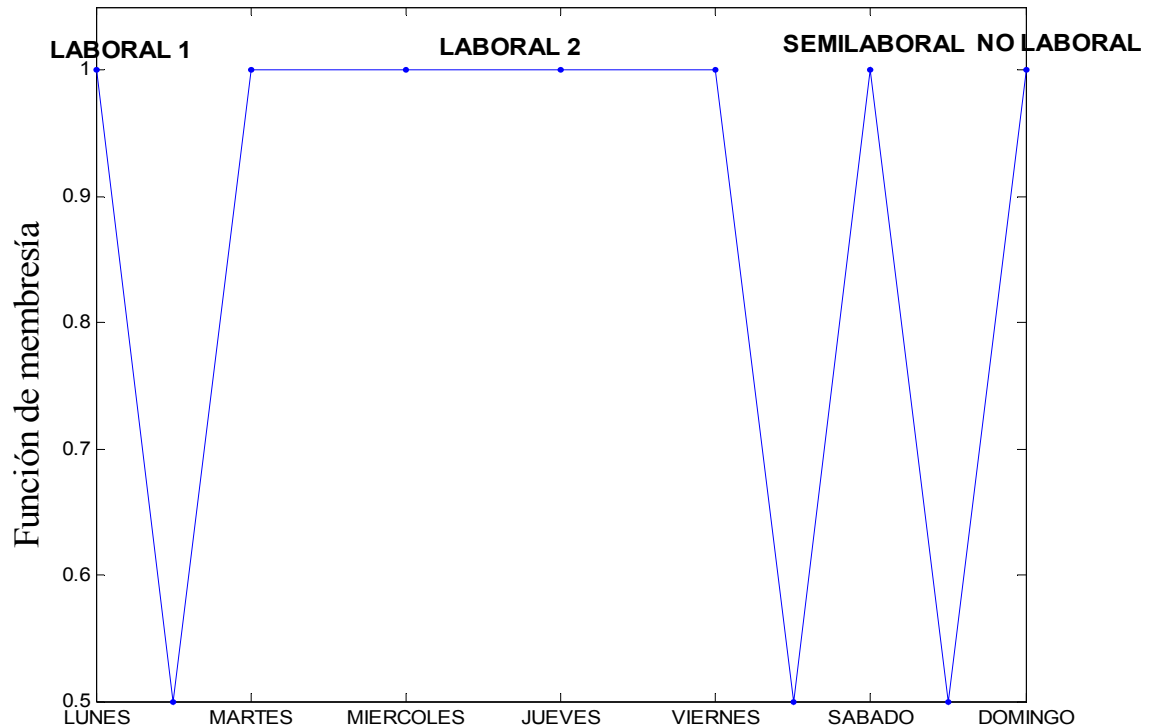


Figura 3.6. Funciones de membresía para la fuzzyficación de la variable Tipo de día

b) Temperatura

La temperatura se fuzzyfica en 2 clases. En la Tabla 3.3 se muestran los límites de las funciones de membresía y la respectiva fuzzyficación de las distintas temperaturas del Sistema ejemplo. Los rangos de variación considerados representativos para las distintas zonas son los que siguen:

- Temperatura mínima de la sierra: de 6 a 10 °C (en valores discretos).
- Temperatura máxima de la sierra: de 14 a 22 °C (en valores discretos).
- Temperatura mínima de la costa: de 18 a 22 °C (en valores discretos).
- Temperatura máxima de la costa: de 25 a 35 °C (en valores discretos).

Tm_sierra	TM_sierra	Tm_costa	TM_costa	clase	codificación	membresía	flanco
[6,8[[14,18[[18,20[[25,27[templado	1	1	0
[8,10]	[18,22]	[20,22]	[27,35]	cálido	2	1	0

Tabla 3.3. Fuzzyficación de la variable Temperatura máxima de la Costa

Las gráficas que se utilizan para la fuzzyficación de la temperatura son funciones trapezoidales, al igual que con la variable “tipo de día”. Para ejemplificar este proceso, se muestra en la Figura 3.7 el caso para la temperatura máxima de la costa (Guayaquil).

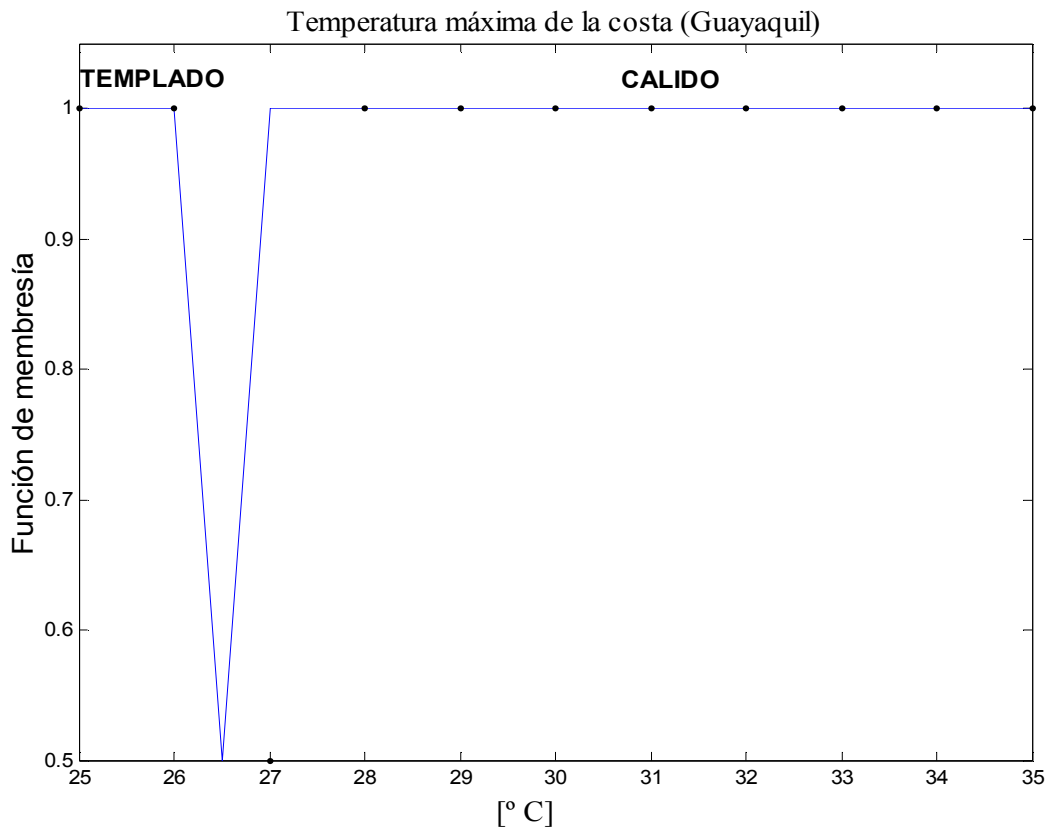


Figura 3.7. Funciones de membresía para la fuzzyficación de la Temperatura máxima de la Costa

c) Demanda

Para modelar la demanda del Sistema ecuatoriano se realiza una división de clases (utilizando la misma división que la variable tipo de día: lunes, martes a viernes, sábado y domingo) como se muestra en la Figura 3.8.

Para una mejor visualización del número de clases para cada tipo de día se muestra el diagrama de duración de carga para la demanda promedio de cada tipo de día, y además se incluye el histograma de las series temporales respectivas desde el 05 de junio hasta el 10 de septiembre de 2000 (base de datos que sirve para obtener la base de conocimiento).

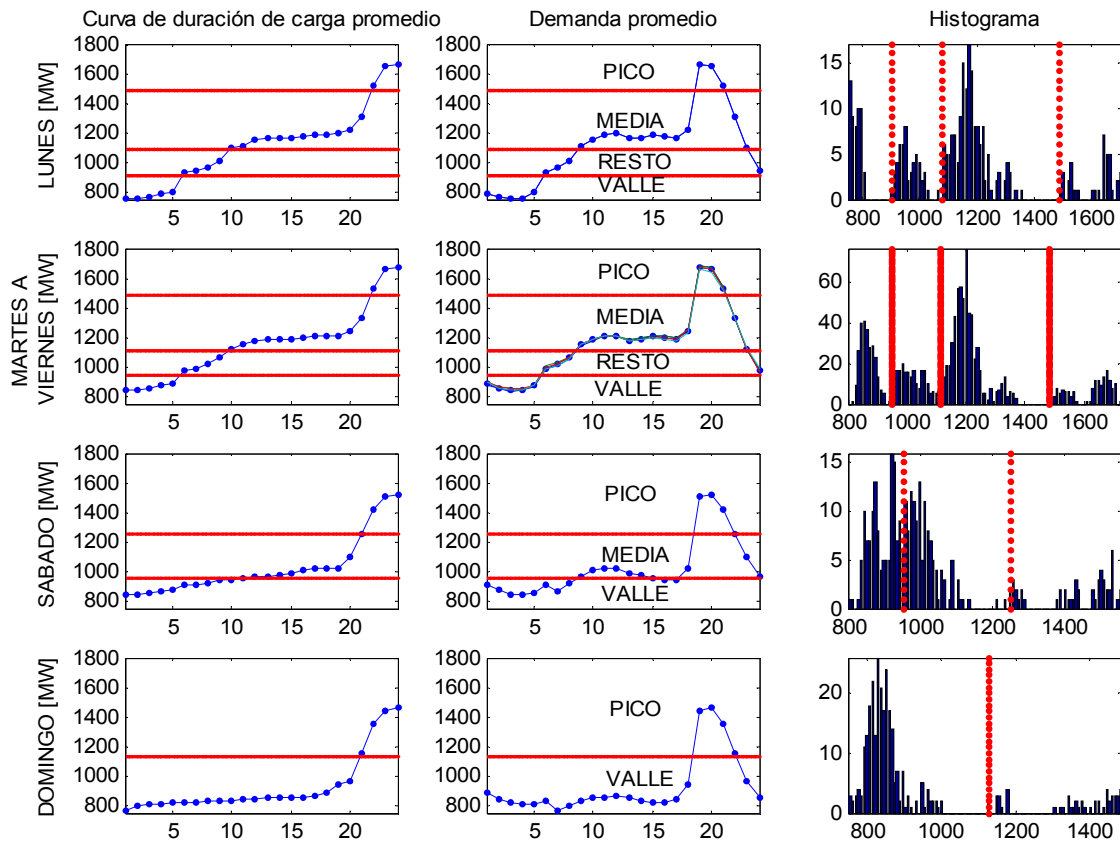


Figura 3.8. Modelación de la Demanda del SSEE ejemplo

La particular distribución de los datos en los histogramas define de una manera “natural”:

- cuatro clases para los días lunes a viernes: pico, media, resto y valle;
- tres clases para el día sábado: pico, media y valle, y
- dos clases para el día domingo: pico y valle.

Además, se incluyen gráficamente las estimaciones de los límites de las funciones de membresía utilizando el algoritmo de Clustering Fuzzy C-means. En la Tabla 3.4 se especifica los límites para cada clase, considerando para la estimación de estos el periodo comprendido entre el 05 de junio hasta el 10 de septiembre de 2000.

clase	codificación	Laboral 1	Laboral 2	Semilaboral	No laboral
valle	1	[747,6;905,1[[793,4;946,4[[791,6;950,8[[746,3;1128,4[
resto	2	[905,1;1083,2[[946,4;1113,8[
media	3	[1083,2;1485,2[[1113,8;1485,4[[950,8;1249,3[
pico	4	[1485,2;1703,6[[1485,4;1729,1[[1249,3;1559,3[[1128,4;1485,2[

Tabla 3.4. Límites de las funciones de membresía para la modelación de la Demanda

Las gráficas que se utilizan para la fuzzyficación de la demanda (funciones gaussianas, ecuación (3.1)), para cada tipo de día (Laboral 1, Laboral 2, Semilaboral y No laboral) se muestran en la Figura 3.9.

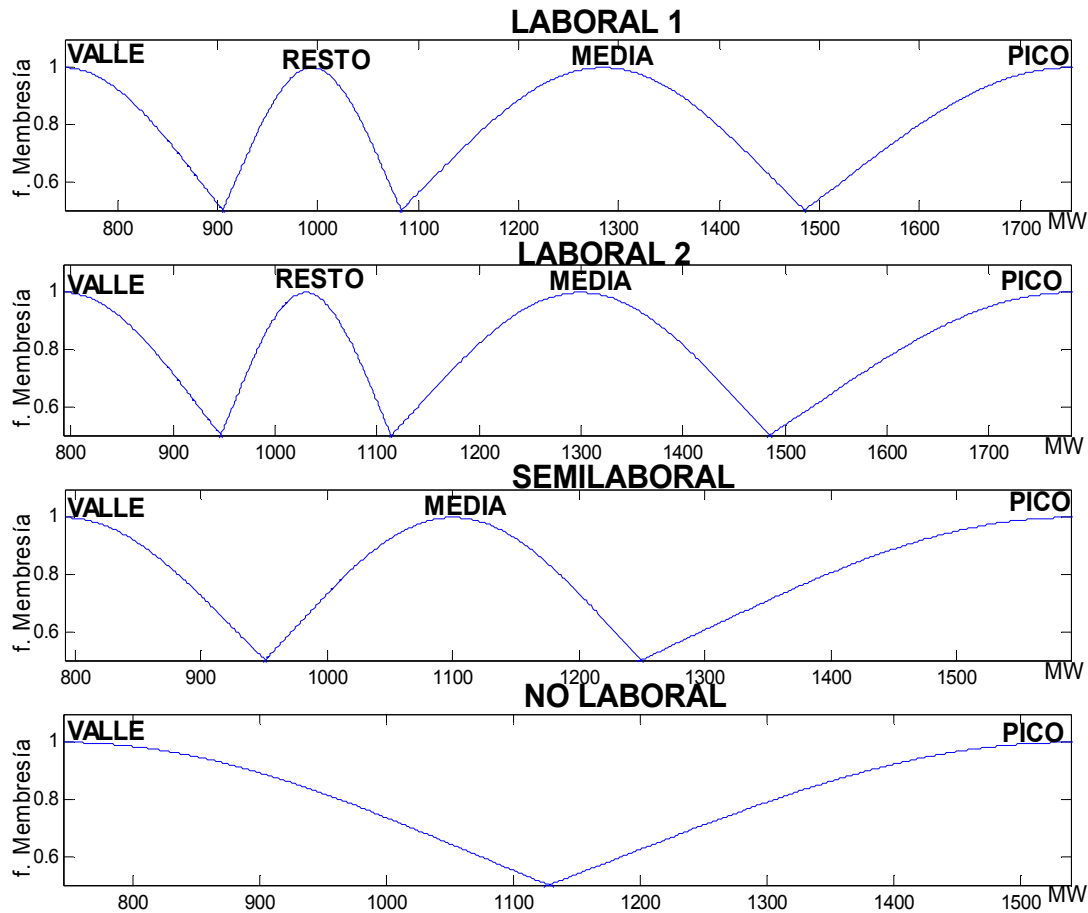


Figura 3.9. Funciones de membresía de “Demanda” según la variable “Tipo de Día”

d) Criterios para los valores máximos de las funciones de membresía

- Debido a que la serie “tipo de día” es determinística, los valores de esta serie siempre están dentro del intervalo [1,7].
- Para la serie “demanda” se considera una holgura cercana a 50 MW, con respecto a los valores máximos de las funciones de membresía de la clase “pico” obtenidos con el algoritmo Fuzzy C-means para cada tipo de día. De esta manera siempre se podrán fuzzyficar los valores de demanda que sean mayores a la demanda pico del periodo considerado, ya que se estima que el máximo incremento futuro de la demanda pico no superará los 50 MW.
Para los valores mínimos de demanda se asume que el incremento natural de la demanda asegura que no existirán valores menores que los ya registrados.

En la Tabla 3.5 se especifican los límites para cada clase, considerando los criterios antes mencionados para la variable “demanda”.

Clase	codificación	Laboral 1	Laboral 2	Semilaboral	No laboral
valle	1	[747,6;905,1[[793,4;946,4[[791,6;950,8[[746,3;1128,4[
resto	2	[905,1;1083,2[[946,4;1113,8[
media	3	[1083,2;1485,2[[1113,8;1485,4[[950,8;1249,3[
pico	4	[1485,2;1750]	[1485,4;1780]	[1249,3;1610]	[1128,4;1535]

Tabla 3.5. Límites máximos de las funciones de membresía para la modelación de la Demanda

3.2.2. Modelación cualitativa

Hasta aquí se ha transformado un sistema multivariable cuantitativo (tipo de día, temperaturas y demanda) en un sistema multivariable cualitativo triple, utilizando el proceso de fuzzyficación descrito en 3.2.1.

En la Figura 3.10 se visualizan a modo de ejemplo 10 registros históricos del sistema multivariable y su representación fuzzy con su correspondiente valor triple cualitativo. En dichas matrices, las letras A, B, C, D, E, F indican las variables del sistema, a saber:

- A: variable tipo de día
- B: variable temperatura mínima de la sierra
- C: variable temperatura máxima de la sierra
- D: variable temperatura mínima de la costa
- E: variable temperatura máxima de la costa
- F: variable demanda

Matriz de códigos de clase del sistema

time	A	B	C	D	E	F
t-2168	1	8	19	20	29	1191.8
t-2167	1	8	19	20	29	1277.3
t-2166	1	8	19	20	29	1695.0
t-2165	1	8	19	20	29	1660.3
t-2164	1	8	19	20	29	1514.4
t-2163	1	8	19	20	29	1316.7
t-2162	1	8	19	20	29	1115.0
t-2161	1	8	19	20	29	953.3
t-2160	2	8	19	21	33	887.9
t-2159	2	8	19	21	33	870.0
...						



Fuzzyficación

Matriz de valores clase

time	A	B	C	D	E	F
t-2168	1	2	2	2	2	3
t-2167	1	2	2	2	2	3
t-2166	1	2	2	2	2	4
t-2165	1	2	2	2	2	4
t-2164	1	2	2	2	2	4
t-2163	1	2	2	2	2	3
t-2162	1	2	2	2	2	3
t-2161	1	2	2	2	2	2
t-2160	2	2	2	2	2	1
t-2159	2	2	2	2	2	1

Matriz de valores de membresía

A	B	C	D	E	F
1	1	1	1	1	0.8637
1	1	1	1	1	0.9992
1	1	1	1	1	0.9705
1	1	1	1	1	0.9235
1	1	1	1	1	0.5777
1	1	1	1	1	0.982
1	1	1	1	1	0.6119
1	1	1	1	1	0.8643
1	1	1	1	1	0.7676
1	1	1	1	1	0.8405

Matriz de valores del flanco

A	B	C	D	E	F
0	0	0	0	0	-1
0	0	0	0	0	-1
0	0	0	0	0	-1
0	0	0	0	0	-1
0	0	0	0	0	-1
0	0	0	0	0	1
0	0	0	0	0	-1
0	0	0	0	0	-1
0	0	0	0	0	1
0	0	0	0	0	1

Matriz triple de valores Cualitativos del sistema

Figura 3.10. Fuzzyficación del sistema multivariable del SSEE ejemplo

En lo que sigue, se realizará la identificación de la máscara óptima de correlación para dicha modelación cualitativa, la cual seleccionará aquellas variables del sistema que tengan una influencia explicativa sobre el PDCP.

Una relación cualitativa para el pronóstico de demanda que incluya las variables “tipo de día”, “temperatura máxima de la Costa” y “demanda histórica”, podría estar dada por la ecuación (3.2), que relaciona el pronóstico de la demanda en (t) con el tipo de día en (t-5), la temperatura máxima de la Costa en (t) y la demanda histórica en 1, 2, 4, 14, 23 y 24 horas anteriores a t.

$$F(t) = \tilde{f}(A(t-5), E(t), F(t-24), F(t-23), F(t-14), F(t-4), F(t-2), F(t-1)) \quad (3.2)$$

donde \tilde{f} es un funcional que denota una relación cualitativa genérica (patrón de entrada-salida) de las variables de entrada.

La ecuación (3.2) puede ser representada por la siguiente máscara de correlación (Tabla 3.6).

t\X	A	B	C	D	E	F
t-24	0	0	0	0	0	-3
t-23	0	0	0	0	0	-4
t-22	0	0	0	0	0	0
t-21	0	0	0	0	0	0
t-20	0	0	0	0	0	0
t-19	0	0	0	0	0	0
t-18	0	0	0	0	0	0
t-17	0	0	0	0	0	0
t-16	0	0	0	0	0	0
t-15	0	0	0	0	0	0
t-14	0	0	0	0	0	-5
t-13	0	0	0	0	0	0
t-12	0	0	0	0	0	0
t-11	0	0	0	0	0	0
t-10	0	0	0	0	0	0
t-9	0	0	0	0	0	0
t-8	0	0	0	0	0	0
t-7	0	0	0	0	0	0
t-6	0	0	0	0	0	0
t-5	-1	0	0	0	0	0
t-4	0	0	0	0	0	-6
t-3	0	0	0	0	0	0
t-2	0	0	0	0	0	-7
t-1	0	0	0	0	0	-8
t	0	0	0	0	-2	+1

Tabla 3.6. Ejemplo de una máscara de correlación

Donde; el eje horizontal **t** muestra la escala temporal y el eje vertical **x** muestra las variables a utilizarse en la máscara de correlación.

El número de filas de la matriz es conocido como la profundidad de la máscara y está directamente relacionada con la observación histórica del sistema. En el ejemplo de la Tabla 3.6 se muestra una máscara de profundidad 25. Los elementos negativos son las *m*-entradas¹ de la máscara (patrones de entrada: i_1, \dots, i_8) y el valor positivo representa la salida (patrón de salida: *o*).

La función de la máscara es seleccionar las *m*-entradas y la salida en cada matriz de clase, membresía y flanco, y transformar dichos valores en patrones estáticos de entrada – salida, o “estados” del sistema. Desplazando la máscara de correlación sobre las matrices de clase, membresía y flanco, se obtienen distintos estados del sistema explicativos del comportamiento del mismo en el tiempo. El conjunto de estados del sistema es conocido como “Matriz de Comportamiento” o “Rule Base” (base de conocimiento). La Figura 3.11 muestra cómo se realiza la construcción de la “matriz de comportamiento”.

¹ Se utiliza *m*-entrada para diferenciar de las entradas del sistema

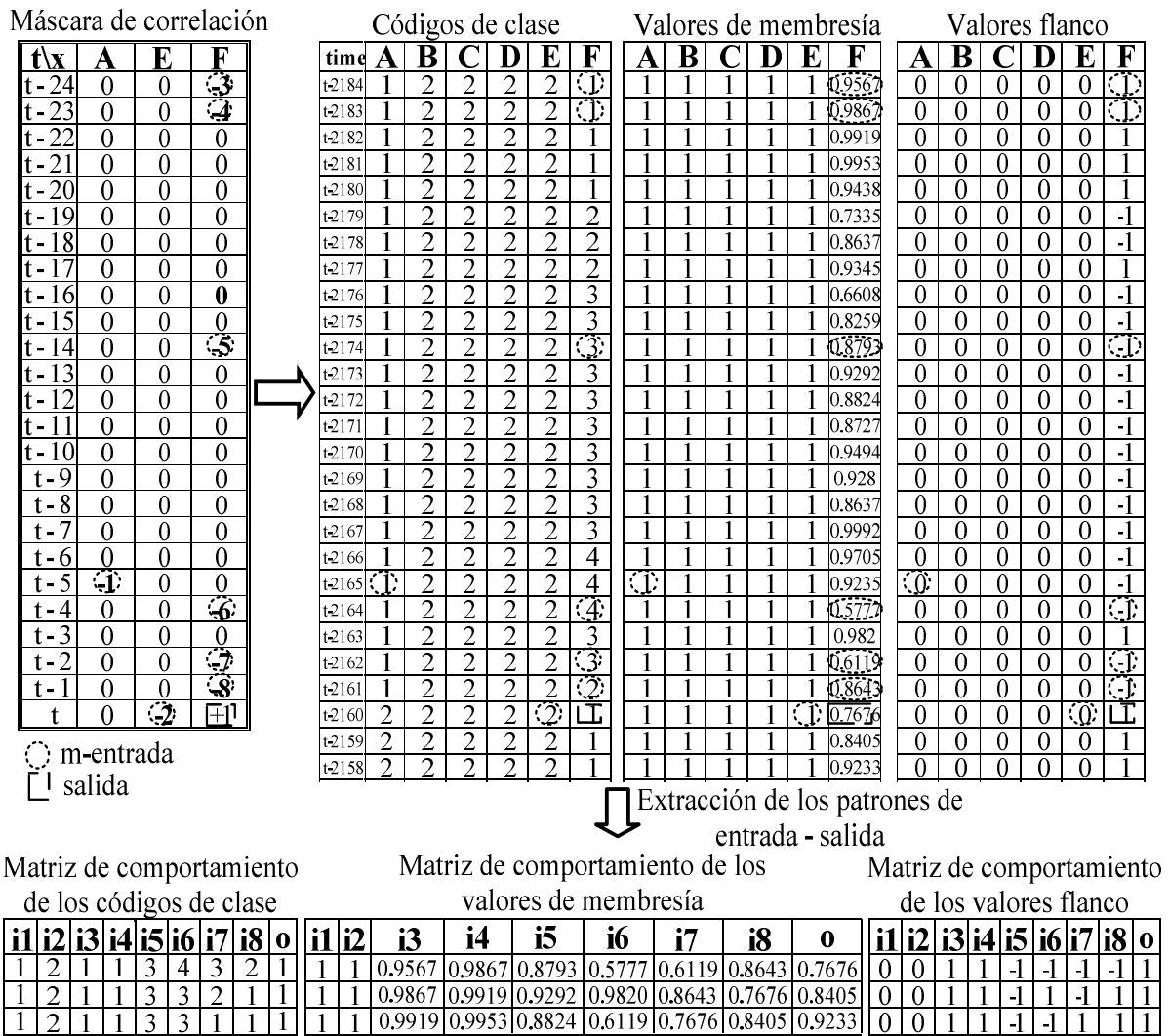


Figura 3.11. Construcción de la matriz de comportamiento del SSEE ejemplo

Si se tienen r registros históricos del sistema a investigar, y considerando que la máscara de correlación posee una profundidad p , entonces es posible obtener $(r - p + 1)$ estados del sistema en la matriz de comportamiento, los cuales conforman la base de conocimiento del sistema.

3.2.3. Simulación cualitativa

En esta etapa es necesario identificar qué estados del sistema registrados en la matriz de comportamiento se asemejan al estado observado al momento de la predicción. De entre aquellos estados registrados más parecidos al estado observado se tomará una predicción de la salida probable del sistema, o pronóstico propiamente dicho.

El esquema se muestra en la Figura 3.12.

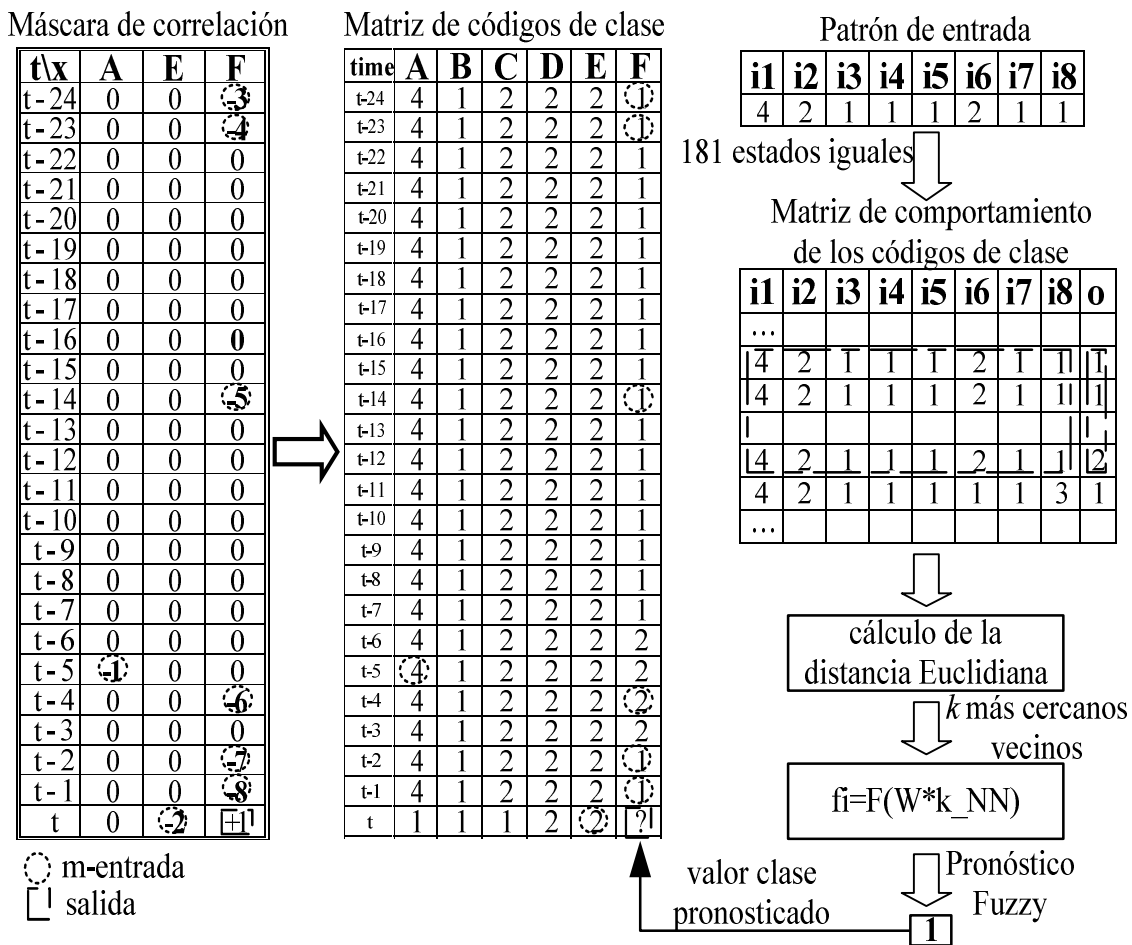


Figura 3.12. Pronóstico fuzzy

El pronóstico opera en dos etapas:

- 1) Etapa determinística: el patrón de entrada para el pronóstico es extraído de la matriz de valores clase por medio de la máscara y es comparado con todos los patrones análogos de la matriz de comportamiento del sistema (datos históricos). Para todos aquellos patrones que coinciden con el patrón de entrada, se observan los valores clase de salida.
- 2) Etapa estadística: se realiza la predicción del valor más probable de salida, a partir de las observaciones anteriores, mediante un promedio ponderado de los valores más próximos en la base de datos de entrenamiento. En esta etapa se utiliza la regla 5-NN (5 más cercanos vecinos).

Las entradas de las variables tipo de día y temperatura sirven para encontrar en la Base de Conocimiento (matriz de comportamiento de los códigos clase) los vecinos más adecuados que se usarán en el proceso de interpolación, es decir son de utilidad para encontrar patrones vecinos medidos durante períodos de tiempo (cronológico y meteorológico) similares.

3.2.4. Defuzzyficación

Las predicciones de los valores de membresía, flanco y código de clase de la salida pronosticada son convertidas a predicciones cuantitativas como se muestra en la Figura 3.13.

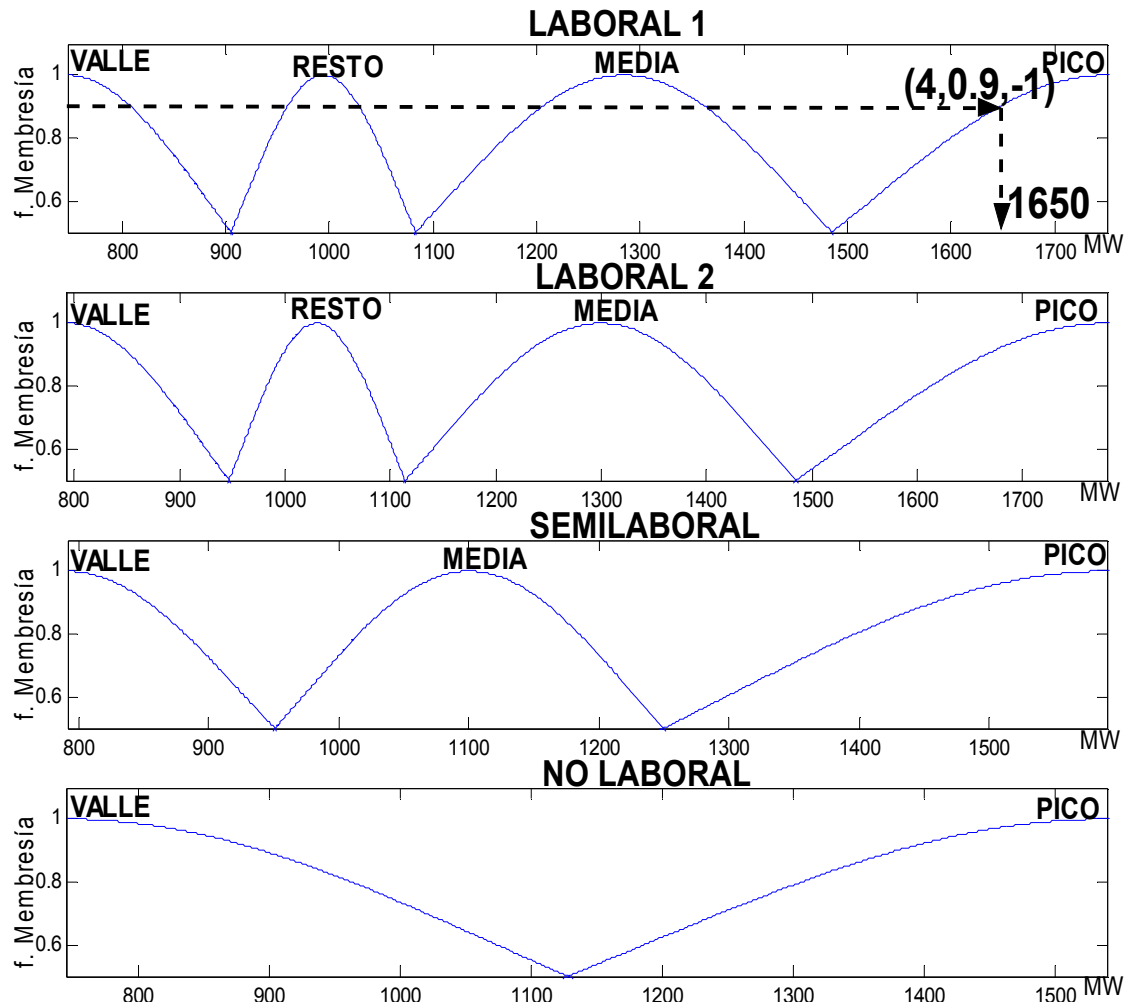


Figura 3.13. Defuzzyficación de la demanda para el día lunes a las 19:00

En la Figura 3.13 se ha supuesto que el estado de entrada observado (4,2,1,1,1,2,1,1) tiene una salida pronosticada igual al triple (4;0,9;-1), el cual –mediante el proceso de defuzzyficación– da un valor para el pronóstico de demanda de 1650 MW.

4. COMPARACIÓN Y SELECCIÓN DE MÁSCARAS DE CORRELACIÓN APROPIADAS

En este capítulo se presenta un método que permite realizar la comparación entre las distintas máscaras de correlación y escoger la que minimice el error del PDCP, utilizando los datos de validación para el cálculo del error relativo.

4.1. ESTRUCTURA DEL PROBLEMA

Como se menciona en el ítem 3.2.2, se tiene que encontrar una máscara de correlación que minimice el error del PDCP. La profundidad de la máscara que se plantea investigar en este ejemplo es de 25 y su estructura de datos se muestra en la Tabla 4.1.

En la referencia [30], se utiliza como herramienta para encontrar la profundidad de la máscara la función de autocorrelación parcial (ACF). En la gráfica de la ACF se puede visualizar la correlación que existe en la serie temporal.

Para el caso de esta investigación, la variable de interés que determinará la profundidad de la máscara será la demanda. En la Figura 2.6 se analiza la ACF de la demanda (05 de junio hasta el 10 de septiembre de 2000). Se grafican 338 retardos, que equivalen a 2 semanas. Si se analiza la gráfica, se determina que la demanda tiene 2 periodicidades: diaria y semanal, para este ejemplo se plantea utilizar para la máscara de correlación una profundidad equivalente a 25 periodos (periodicidad diaria). Máscaras muy profundas (169 períodos) no serán consideradas en esta investigación debido a que se las considera computacionalmente ineficientes.

Si se decide trabajar con máscaras profundas, por ejemplo la de periodicidad semanal, Cellier propone un método para limitar la búsqueda. Primeramente, se plantea una máscara de profundidad pequeña (diaria) para ver cuales son las entradas usadas por las mejores máscaras de esa profundidad. Se aumenta la profundidad de la máscara considerando las entradas encontradas anteriormente (a las entradas que no tienen ninguna correlación, encontradas en el proceso anterior, no se las considera en la máscara candidata) y se localiza más entradas que pudieran ser útiles en la modelación. Se continúa de esta forma hasta que se obtiene la profundidad deseada.

En el capítulo 5 se analizará más detalladamente la determinación de la profundidad de la máscara a utilizar en el PDCP.

t\X	A	B	C	D	E	F
t-24	1	26	51	76	101	126
t-23	2	27	52	77	102	127
t-22	3	28	53	78	103	128
t-21	4	29	54	79	104	129
t-20	5	30	55	80	105	130
t-19	6	31	56	81	106	131
t-18	7	32	57	82	107	132
t-17	8	33	58	83	108	133
t-16	9	34	59	84	109	134
t-15	10	35	60	85	110	135
t-14	11	36	61	86	111	136
t-13	12	37	62	87	112	137
t-12	13	38	63	88	113	138
t-11	14	39	64	89	114	139
t-10	15	40	65	90	115	140
t-9	16	41	66	91	116	141
t-8	17	42	67	92	117	142
t-7	18	43	68	93	118	143
t-6	19	44	69	94	119	144
t-5	20	45	70	95	120	145
t-4	21	46	71	96	121	146
t-3	22	47	72	97	122	147
t-2	23	48	73	98	123	148
t-1	24	49	74	99	124	149
t	25	50	75	100	125	150

Tabla 4.1. Estructura de datos para la máscara de correlación (Matriz M)

En la Tabla 4.1. cada fila muestra en horas los registros históricos ($t-1, \dots, t-25$: “profundidad de la máscara”) y cada columna muestra las variables a utilizar en la máscara de correlación, donde:

- A: variable tipo de día
- B: variable temperatura mínima de la sierra
- C: variable temperatura máxima de la sierra
- D: variable temperatura mínima de la costa
- E: variable temperatura máxima de la costa
- F: variable demanda

La matriz de datos M que de la Tabla 4.1. surge, es de dimensión 25x6. De los 150 elementos de esta matriz, 149 son datos factibles (inputs) y el elemento $M_{25,6}$ -que es el valor de demanda a pronosticar en el instante t – es el resultado (output).

Considerando como función objetivo al error del pronóstico de demanda de corto plazo, el proceso de optimización debe conducir a definir una máscara con aquellos elementos correspondientes a datos (inputs o entradas) de la matriz M tales que, en la metodología FIR, el resultado (output o salida) minimice esta función objetivo.

Como ejemplo, si se toma el caso especial de una máscara que considere *una entrada para cada variable* ($m=6$), el problema de optimización debe formularse de la siguiente manera:

Minimizar:	$\text{Error_PDCP}(a_1, b_1, c_1, d_1, e_1, f_1)$
Tal que:	$1 \leq a_1, b_1, c_1, d_1, e_1 \leq 25 \quad \& \quad 1 \leq f_1 \leq 24$
donde;	a_1 : elemento de la variable tipo de día b_1 : elemento de la variable temperatura mínima de la sierra c_1 : elemento de la variable temperatura máxima de la sierra d_1 : elemento de la variable temperatura mínima de la costa e_1 : elemento de la variable temperatura máxima de la costa f_1 : elemento de la variable demanda

4.2. ESPACIO DE SOLUCIONES DEL PDCP

El espacio de soluciones del PDCP está compuesto por todas las configuraciones factibles de la máscara candidata (CFMC). Las CFMC pertenecen al conjunto Θ_{CFMC} del universo $|\Theta|$ formado por todas las combinaciones posibles de las m -entradas.

Definimos además el conjunto Θ_N que esta formado por todas las configuraciones que no son factibles, es decir aquellas configuraciones que arman patrones de entrada que no poseen un patrón similar o comparable en la matriz de comportamiento, donde $|\Theta| = \Theta_{\text{CFMC}} \cup \Theta_N$. Dentro del conjunto Θ_{CFMC} es necesario identificar la máscara de correlación óptima del PDCP.

4.2.1. Formulación simbólica de una configuración factible (CFMC)

Una determinada configuración ν de la máscara de correlación puede definirse mediante la posición que toma cada una de las m -entradas en la máscara. Matemáticamente, una configuración es el conjunto ν definido por la ecuación (4.1).

$$\nu = (x_1, x_2, \dots, x_m) \forall m, \nu \in \Theta \in \mathfrak{S}^m \tag{4.1}$$

donde; m : número de entradas de la máscara candidata.
 x_i : entero que representa la posición de la m -entrada en la máscara ν y $1 \leq x_i \leq 149$.

Las posiciones de las entradas en la máscara de correlación se interpretan como una relación cualitativa. Por ejemplo, para la relación cualitativa de la ecuación (3.2), la configuración equivalente es $\nu=(20,125,126,127,128,134,136,146,148,149)$.

El problema de optimización se plantea entonces como:

$$\text{Error_PDCP}(\mathbf{v}^*) \rightarrow \min.$$

$$\text{Tal que: } \mathbf{v}^* \in \Theta_{CFMC}$$

4.2.2. Dimensión del problema

A partir de las consideraciones planteadas en el problema de optimización de tipo combinatorial, y considerando la matriz M de correlación de la Tabla 4.1, es posible determinar el número total de configuraciones posibles, equivalente al tamaño del espacio de soluciones $|\Theta|$, utilizando la ecuación (4.2).

$$|\Theta| = \binom{149}{m} = \frac{149!}{(149-m)!m!} \quad (4.2)$$

De la ecuación (4.2), para $m = 6$ tenemos que $|\Theta| = 13.725.120.696$, es decir, más de 13 mil millones de combinaciones *posibles* de las m -entradas.

Para el problema de optimización planteado como ejemplo en el ítem 4.1 (1 entrada para cada variable), las configuraciones *factibles* de la máscara candidata se reducen a:

$$|\Theta_{CFMC}|(\mathbf{v}_I) = \prod_m X_m$$

donde; X_m : posiciones a variar en la máscara para cada m -entrada (variable).
 $\mathbf{v}_I = (a_1, 25+b_1, 50+c_1, 75+d_1, 100+e_1, 125+f_1)$

y el espacio de configuraciones factibles $|\Theta_{CFMC}|$ para el problema de optimización planteado en el ítem 4.1 da como resultado:

$$|\Theta_{CFMC}|(\mathbf{v}_I) = 25*25*25*25*25*24 = 234.375.000$$

Esto significa que, aún para el ejemplo considerado, se cuenta con más de 234 millones de configuraciones posibles de ser evaluadas con la función objetivo (Error del PDCP).

Considerando otro ejemplo, en el cual la máscara contiene dos datos de la variable demanda, la configuración de la solución \mathbf{v} se amplía en una dimensión:

$$\mathbf{v}_{II} = (a_1, 25+b_1, 50+c_1, 75+d_1, 100+e_1, 125+f_1, 125+f_2),$$

donde f_2 es un elemento de la variable “demanda” tal que $1 \leq f_2 \leq 24$ y $f_2 \neq f_1$. El espacio de soluciones para esta configuración viene dado por la ecuación (4.3).

$$|\Theta| = \prod_m X_m * \binom{j}{k} \quad \forall m \neq j \quad \& \quad (m=1, \dots, 6) \quad (4.3)$$

donde;
$$\binom{j}{k} = \frac{j*(j-1)*(j-1)*...*(j-k+1)}{k!} = \frac{j!}{k!*(j-k)!}$$

j : posiciones a variar

k : número de entradas para una misma variable

La ecuación (4.3) se emplea para calcular el número de combinaciones factibles cuando se tiene más de 1 entrada para una misma variable. Para la configuración v_{11} (2 entradas para la variable demanda) el número de posibles soluciones será:

$$|\Theta_{CFMC}|(v_{11}) = 25*25*25*25*25*[24!/(2!*22!)] = 2.695.300.000.$$

4.2.3. Número de configuraciones a ser evaluadas por el SRA

Con el objeto de disminuir el espacio de soluciones factibles y asumiendo que las entradas de temperatura correspondientes a las temperaturas máxima y mínima de la Sierra y mínima de la Costa no poseen un grado de correlación con la demanda que sea significativo, y por tanto no aportan mayor información al PDCP [43], se plantea la eliminación de dichas variables exógenas para la formulación del problema de optimización. Por otro lado, es esperable que la demanda histórica pueda dar mayor información para el pronóstico – recordemos que los modelos econométricos se basan exclusivamente en estos datos – por lo cual será deseable incorporar más entradas en la máscara de correlación que consideren a la variable “demanda”.

En la Tabla 4.2 se detalla la dimensión del espacio de soluciones utilizando la ecuación (4.3) y variando entre 1 y 8 la cantidad de entradas para la variable “demanda” ($|\Theta|_D$), pero considerando solamente 1 entrada en la variable “tipo de día” y 1 entrada para la variable “temperatura máxima de la Costa”, manteniendo además una máscara de correlación de profundidad 25.

v	f_i		$ \Theta _D$	$ \Theta =25*25* \Theta _D$
v_3	f_1	$24!/(1!*23!)$	24	15.000
v_4	f_1, f_2	$24!/(2!*22!)$	276	172.500
v_5	f_1, f_2, f_3	$24!/(3!*21!)$	2.024	1.265.000
v_6	f_1, \dots, f_4	$24!/(4!*20!)$	10.626	6.641.250
v_7	f_1, \dots, f_5	$24!/(5!*19!)$	42.504	26.565.000
v_8	f_1, \dots, f_6	$24!/(6!*18!)$	134.596	84.122.500
v_9	f_1, \dots, f_7	$24!/(7!*17!)$	346.104	216.315.000
v_{10}	f_1, \dots, f_8	$24!/(8!*16!)$	735.471	459.669.375

Tabla 4.2. Espacio de soluciones para la máscara de la Tabla 3.6

4.2.4. Función de evaluación de una configuración CFMC

El método de entrenamiento que se propone aplicar en esta investigación es el “**Aprendizaje Supervisado**”, utilizando como valor de referencia los datos horarios de la demanda de la última semana de entrenamiento. Para encontrar la configuración \mathbf{v}^* , tal que esta configuración minimice la sumatoria de los errores obtenidos en la última semana de entrenamiento, se utiliza -para el cálculo del error- el pronóstico de demanda para 1 hora en adelante² y la demanda real histórica para dicha hora.

El análisis del espacio de soluciones se realiza mediante la minimización de la función de evaluación o “función objetivo”, que es igual al promedio de los últimos 168 errores relativos del pronóstico horario respecto del valor real histórico de la demanda para la semana de entrenamiento, según lo expresa la ecuación (4.4).

$$Error_PDCP(\mathbf{v}) = E_{(t)} = \frac{\sum_{i=1}^{168} \left(\frac{|\hat{D}_{t-i} - D_{real,t-i}|}{D_{real,t-i}} \right)}{168} * 100 [\%] \quad (4.4)$$

donde; \hat{D}_t es el PDCP para la hora t y $D_{real,t}$ es la demanda histórica real de la hora t durante la última semana de los datos de entrenamiento.

4.2.5. Datos de entrenamiento y datos de validación o test

Se denomina “datos de entrenamiento” a los datos de demanda histórica utilizados para el cálculo de la función objetivo planteada en (4.4) y al conjunto de datos de las variables de entrada utilizados para el cálculo del pronóstico (D_{t-i}) de cada hora $t-i$ necesario para la determinación de $E_{(t)}$. La configuración óptima \mathbf{v}^* es aquella que minimiza el error $E_{(t)}$ utilizando los datos de entrenamiento:

$$Error_PDCP(\mathbf{v}^*) = E_{(t)} \rightarrow \min.$$

Se denominan “datos de validación” o “datos de test” al conjunto de datos que se utilizan para asegurar la validez de la configuración óptima \mathbf{v}^* encontrada para la hora t . Utilizando la máscara de correlación dada por la configuración \mathbf{v}^* , se realiza el PDCP para las próximas horas, es decir para: $D_{(t+i)}$, con $i > 0$. En este proceso, para el PDCP se asumen conocidos (reales) todos los datos anteriores a las próximas 24 horas a pronosticar (1 día en adelante), por lo cual todos estos datos se toman de la base de datos existente. De esta manera, con los datos de validación se calcula el error del pronóstico de demanda de un día en adelante efectivamente producido por el método FIR.

Los datos de validación o test posibilitan evaluar el alcance de la validez de la máscara de correlación. En efecto, si el error $E_{TEST(t)}$ del PDCP calculado con los “datos de test” según la ecuación (4.5) resulta tolerable para $i=1, \dots, 168$, entonces la máscara de

² “1 hora de adelante” significa que se asumen conocidos todos los datos para las horas anteriores a la pronosticada. “1 día de adelante” significa que se pronostican 24 horas para $t+i$ ($i=0, \dots, 23$), considerando conocidos (reales) todos los datos para las horas $t-j$ ($j > 0$) y tomando los valores pronosticados ya calculados entre la hora t y $t+i-1$.

correlación correspondiente a la configuración \mathbf{v}^* podrá considerarse válida para las próximas 168 horas, posibilitando así el pronóstico de demanda horaria de un día completo en adelante durante toda una semana. En caso contrario, la máscara de correlación deberá ser recalculada con mayor frecuencia.

$$E_{TEST(t)} = \frac{\sum_{i=1}^{168} \left(\frac{|\hat{D}_{t+i} - D_{real,t+i}|}{D_{real,t+i}} \right)}{168} * 100 [\%] \quad (4.5)$$

donde; \hat{D}_t es el PDCP utilizando \mathbf{v}^* para la hora t con hasta un día de adelanto y $D_{real,t}$ es la demanda histórica real de la hora t durante la semana de validación o test.

Debido a que la máscara óptima de correlación se encuentra mediante la optimización de la función objetivo que considera 1 semana de entrenamiento y tiene una validez para el PDCP de 7 días en adelante, se podría con esta máscara realizar el pronóstico de demanda para 1 semana de adelanto, el cual se usa para la Programación de los embalses semanales.

El análisis de la validez de la máscara de correlación será analizado con mayor detenimiento en el capítulo 5.

4.3. ALGORITMOS EVOLUTIVOS PARA LA DETERMINACIÓN DE LA CFMC ÓPTIMA

Una primera aproximación para la búsqueda de la configuración óptima es la evaluación exhaustiva del espacio de soluciones, pero debido al tamaño de posibles combinaciones -como se muestra en la Tabla 4.2- y al elevado requerimiento en tiempo de cálculo que solicita el PDCP, se concluye que no sería posible realizar una evaluación exhaustiva del espacio de soluciones, por lo que se plantea el uso de algoritmos evolutivos para la solución del Problema de Optimización.

4.3.1. Minimizar la función objetivo mediante búsqueda local aleatoria

El problema de optimización planteado hasta el momento posee las siguientes propiedades:

- es un problema de naturaleza combinatorial (NP-Hard).
- con un dominio no convexo.
- con una función objetivo no lineal y no analítica.

En la referencia [22] se determinó que tales funciones poseen un número considerable de óptimos locales. Como consecuencia de esto, la determinación del óptimo global es una ardua tarea en lo que respecta a elevados tiempos de cálculo y algoritmos de optimización complicados.

Para la resolución del problema de optimización resultan apropiados los **Algoritmos Evolutivos** (AE) [1, 14, 17, 23 y 33], debido a la calidad de la solución que encuentran y a los tiempos de cálculo relativamente reducidos (en comparación con la búsqueda exhaustiva) que emplean para alcanzarla. Además de esto, los AE posibilitan la determinación de varias soluciones subóptimas cercanas al óptimo global.

Los AE pertenecen a la clase de algoritmos que realizan el proceso de optimización mediante búsqueda local aleatoria. Puesto que son Algoritmos de Aproximación [1], no necesariamente alcanzan el óptimo global. Sin embargo, y dependiendo de la performance de cada algoritmo, las soluciones finales son muy cercanas al óptimo global cuando este no alcanza a ser determinado.

Para la minimización de la función objetivo se consideró la aplicación del algoritmo **Simulated Rebounding Algorithm** (SRA), que pertenece a la clase de algoritmos estocásticos (randomization algorithms) basados en búsqueda local aleatoria controlada (como lo es, entre otros, el Simulated Annealing). El SRA ha sido contrastado con el Simulated Annealing y algunas implementaciones de Algoritmos Genéticos en problemas de optimización con similares características [45] y debido a la alta performance del mismo es que se elige en esta investigación el SRA como motor de optimización del problema aquí planteado (ver apéndice C).

Mediante la aplicación del SRA es posible determinar un conjunto de soluciones subóptimas muy cercanas al óptimo global. Se ha estimado que el tiempo de cálculo que emplea el SRA es menor que el 5% del necesario para la evaluación exhaustiva del espacio de soluciones factibles [22 y 45].

4.3.2. Simulated Rebounding Algorithm (SRA)

Para la implementación del SRA es necesario:

- Una condición o configuración de inicio (v_o).
- Una estructura de vecindad V (definición de configuraciones cercanas o adyacentes a la evaluada).
- Un procedimiento para la construcción de próximas configuraciones a ser evaluadas.
- El valor del parámetro de simulación α (coeficiente de elasticidad).

El SRA analiza solamente una porción del espacio de soluciones factibles y no garantiza la determinación del óptimo global. La cantidad de los subóptimos encontrados depende de los parámetros que se definen y por lo tanto es recomendable la realización del proceso de optimización partiendo de distintas condiciones iniciales. Una aplicación múltiple tiene además el fin de determinar la mayor cantidad posible de óptimos locales cercanos al global [45].

Puesto que el conjunto de soluciones subóptimas se encuentra analizando la trayectoria seguida por el algoritmo para encontrar la solución óptima (espacio de soluciones evaluado) y con el fin de que este conjunto sea lo más representativo del espacio de solución total $|\Theta|$, normalmente se consideran un número de soluciones iniciales v_o igual a la dimensión del espacio Θ (número de parámetros a optimizar, en nuestro caso el número m de entradas de la máscara de correlación).

En la Tabla 4.3, se muestra un ejemplo de las condiciones iniciales que deben ser evaluadas para resolver el problema de optimización para la configuración v_8 (Tabla 4.2).

Condiciones Iniciales
$v_8^1=(25,125,144,145,146,147,148,149)$
$v_8^2=(25,125,126,145,146,147,148,149)$
$v_8^3=(25,125,126,127,146,147,148,149)$
$v_8^4=(25,125,126,127,128,147,148,149)$
$v_8^5=(25,125,126,127,128,129,148,149)$
$v_8^6=(25,125,126,127,128,129,130,149)$
$v_8^7=(25,125,126,127,128,129,130,131)$

Tabla 4.3. Condiciones iniciales para la configuración v_8

Estas condiciones iniciales pueden considerarse como los vértices más alejados pertenecientes al hiperespacio de soluciones $|\Theta|$.

La estructura de vecindad Ω_v define el conjunto de configuraciones v_k más cercanas a v , y puede definirse matemáticamente para esta instancia específica con la ecuación (4.5).

$$\Omega_v = v_k \Leftrightarrow \begin{cases} x_{j(v_k)} = x_{j(v)} \pm n & (j = 1, \dots, m) \\ x_{i(v_k)} = x_{i(v)} & (i \neq j) \end{cases} \quad (4.5)$$

donde; m : es el número de entradas de la máscara de correlación
 n : número positivo aleatorio con distribución uniforme

Se ha considerado que para un profundidad de 25, el valor máximo de n es de 6. Valores mayores ocasionarían un comportamiento caótico y no se logrará una búsqueda eficiente en el espacio de soluciones planteado.

El procedimiento a seguir para la construcción de una solución vecina para la configuración v_8 , considerando $m = 8$ y que la estructura de vecindad podrá variar con $n \leq 6$, será:

- elegir un j entre 1 y 8 en forma aleatoria con probabilidad uniforme,
- elegir aleatoriamente con probabilidad uniforme un número entre ± 6 y
- variar el $x_{j(v)}$ según el valor obtenido anteriormente.

El tamaño de tal estructura de vecindad será:

$$|\Theta_\Omega| = 8 * 12 = 96.$$

El valor del parámetro α varía según [28] entre $0,7 < \alpha < 0,99$.

El SRA se lanza con las diferentes condiciones iniciales v_8 , y para cada variante se asigna un valor $c_{(v)}$ a fin de identificar aquellas ya evaluadas durante el proceso.

$$c_{(v_i)} = \begin{cases} 0 & \text{para } v_i \text{ aún no simulada} \\ < 0 & \text{para } v_i \text{ no factible } (\Theta_N) \\ E & \text{para } v_i \text{ factible } (\Theta_{CFMC}) \end{cases}$$

Luego de correr el algoritmo SRA para las 7 condiciones iniciales de la Tabla 4.3, se deberá elegir la configuración factible v^* que obtenga el menor error $E_{(t)}$ del PDCP para la semana de entrenamiento (mínimo minimorum) y esta configuración CFMC_{óptima} será la solución del problema de optimización planteado.

A continuación se muestra un ejemplo de la aplicación del algoritmo SRA, para la configuración v_8 y para el periodo de test de la semana del 11 al 17 de septiembre, utilizando para el entrenamiento la semana del 04 al 10 de septiembre de 2000, corriendo cada condición inicial 4 veces.

C_Ini 1: {A(t),E(t),F(t-6),F(t-5),F(t-4),F(t-3),F(t-2),F(t-1)} $E_{(t)}=1,8157$ & $E_{TEST(t)}=1,4914$

soluciones encontradas	$E_{(t)}$	$E_{TEST(t)}$
A(t-5),E(t-3),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1943	1,0443
A(t-5),E(t-5),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1943	1,0429
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-23),F(t-17),F(t-10),F(t-4),F(t-2),F(t-1)	1,2357	1,1771
A(t-4),E(t-2),F(t-10),F(t-5),F(t-4),F(t-3),F(t-2),F(t-1)	1,6229	1,2543

C_Ini 2: {A(t),E(t),F(t-24),F(t-5),F(t-4),F(t-3),F(t-2),F(t-1)} $E_{(t)}=1,6271$ & $E_{TEST(t)}=1,3143$

soluciones encontradas	$E_{(t)}$	$E_{TEST(t)}$
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t-14),F(t-24),F(t-18),F(t-11),F(t-5),F(t-2),F(t-1)	1,2586	1,2114
A(t-5),E(t),F(t-24),F(t-23),F(t-13),F(t-5),F(t-2),F(t-1)	1,2100	1,1257
A(t-5),E(t-5),F(t-23),F(t-13),F(t-10),F(t-5),F(t-2),F(t-1)	1,2200	1,0614
A(t-5),E(t-13),F(t-24),F(t-17),F(t-15),F(t-12),F(t-2),F(t-1)	1,1971	1,0314

C_Ini 3: {A(t),E(t),F(t-24),F(t-23),F(t-4),F(t-3),F(t-2),F(t-1)} $E_{(t)}=1,5286$ & $E_{TEST(t)}=1,1129$

soluciones encontradas	$E_{(t)}$	$E_{TEST(t)}$
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800

C_Ini 4: {A(t),E(t),F(t-24),F(t-23),F(t-22), F(t-3),F(t-2),F(t-1)} $E_{(t)}=1,3286$ & $E_{TEST(t)}=1,2143$

soluciones encontradas	$E_{(t)}$	$E_{TEST(t)}$
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t-13),F(t-24),F(t-17),F(t-15),F(t-12),F(t-2),F(t-1)	1,1971	1,0314
A(t-2),E(t),F(t-24),F(t-23),F(t-18),F(t-16),F(t-2),F(t-1)	1,1943	1,1443
A(t-5),E(t-21),F(t-24),F(t-23),F(t-16),F(t-15),F(t-10),F(t-1)	1,2414	1,1871

C_Ini 5: {A(t),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-2),F(t-1)} $E_{(t)}=1,3914$ & $E_{TEST(t)}=1,1814$

soluciones encontradas	$E_{(t)}$	$E_{TEST(t)}$
A(t-5),E(t-3),F(t-24),F(t-23),F(t-22),F(t-18),F(t-2),F(t-1)	1,1771	1,0543
A(t-5),E(t-17),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,1957	1,0514
A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,1657	0,9600
A(t-4),E(t-5),F(t-24),F(t-23),F(t-22),F(t-12),F(t-2),F(t-1)	1,2100	1,0557
A(t-5),E(t-6),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,2071	1,0171

C_Ini 6: {A(t),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-20),F(t-1)} $E_{(t)}=1,4186$ & $E_{TEST(t)}=1,1557$

soluciones encontradas	$E_{(t)}$	$E_{TEST(t)}$
A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,1657	0,9600
A(t-3),E(t-10),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,2086	1,0571
A(t-5),E(t-6),F(t-24),F(t-23),F(t-22),F(t-19),F(t-2),F(t-1)	1,2486	1,0843
A(t-5),E(t),F(t-24),F(t-23),F(t-17),F(t-16),F(t-15),F(t-1)	1,2500	1,0700
A(t-3),E(t-10),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,2086	1,0571

C_Ini 7: {A(t),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-20),F(t-19)} $E_{(t)}=1,3329$ & $E_{TEST(t)}=1,0929$

soluciones encontradas	$E_{(t)}$	$E_{TEST(t)}$
A(t-5),E(t-4),F(t-24),F(t-23),F(t-22),F(t-16),F(t-14),F(t-1)	1,1886	0,9986
A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-2),F(t-1)	1,2314	1,0114
A(t-3),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-2),F(t-1)	1,2200	0,9071
A(t-4),E(t-5),F(t-24),F(t-23),F(t-22),F(t-17),F(t-2),F(t-1)	1,2129	0,9686
A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-2),F(t-1)	1,1729	0,9829

Tiempo de cálculo = 13,48 minutos.

En la Tabla 4.4 se muestra el resumen de las máscaras de correlación con menor error $E_{(t)}$ de la semana de entrenamiento para cada una de las siete condiciones iniciales simuladas por el algoritmo FIR & SRA. Se indican los respectivos errores relativos promedios de entrenamiento y test.

$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$	$E_{(t)}$	$E_{TEST(t)}$
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	1,1929	1,0800
A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,1657	0,9600
A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,1657	0,9600
A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-2),F(t-1)	1,1729	0,9829

Tabla 4.4. Resumen de las soluciones encontradas por el FIR & SRA para cada condición inicial

La solución óptima que obtiene el menor error $E_{(t)}$ del PDCP para la semana de entrenamiento (mínimo minimorum) se encuentra con las condiciones iniciales 5 y 6,

dando como resultado un error $E_{(t)}$ igual a 1,1657%. Con las m -entradas obtenidas a través de la optimización de la función objetivo planteada en la ecuación (4.4), se realiza el PDCP para la semana de test del 11 al 17 de septiembre de 2000 y se obtiene un error de test promedio $E_{TEST(t)}$ igual a 0,9600%.

5. EVALUACIÓN DE LAS MEJORES MÁSCARAS DE CORRELACIÓN UTILIZANDO FIR & SRA

En este capítulo se muestran los resultados que arroja el algoritmo evolutivo SRA y los criterios de selección que permiten determinar, de entre todas las mejores máscaras de correlación, cuál es la óptima para el PDCP.

La aplicación y análisis de FIR al PDCP se realiza en forma detallada para días “normales” de la semana, excluyendo las atipicidades propias de los días “feriados”.

No se incluye en esta Investigación los días “feriados” debido a que estos días para el SSEE del Ecuador son pocos en el año y la curva de demanda año a año para cada uno de estos días tiene las mismas características, por lo que para el pronóstico su curva típica adimensional se escala de acuerdo al crecimiento de demanda del sistema. Por lo tanto no es necesario incorporarlos a la Metodología de solución propuesta.

5.1. PDCP PARA DÍAS NORMALES (NO FERIADOS)

5.1.1. Datos del Sistema Ejemplo

La base de datos del sistema multivariable (Tabla 3.1) a utilizar en esta investigación, consiste en registros horarios de temperatura y demanda para el periodo desde el 12 de junio hasta el 24 de septiembre de 2000. Los grupos de datos de entrenamiento y de test se muestran en la Tabla 5.1. La estructura de la base de datos se muestra en la Tabla 5.2.

Mes	lu	ma	mi	ju	vi	sa	do
junio	12	13	14	15	16	17	18
	19	20	21	22	23	24	25
	26	27	28	29	30		
julio						1	2
	3	4	5	6	7	8	9
	10	11	12	13	14	15	16
	17	18	19	20	21	22	23
	24	25	26	27	28	29	30
	31						
agosto		1	2	3	4	5	6
	7	8	9	10	11	12	13
	14	15	16	17	18	19	20
	21	22	23	24	25	26	27
	28	29	30	31			
septiembre					1	2	3
	4	5	6	7	8	9	10
	11	12	13	14	15	16	17
	18	19	20	21	22	23	24

Base de Datos de validación o test

Base de Datos de entrenamiento

Semana de entrenamiento

Semanas de test

Tabla 5.1. Base de datos del Sistema Ejemplo

Tipo de día	T_min sierra	T_max sierra	T_min costa	T_max costa	Demanda
day (t-2184)	Tm_Uio (t-2184)	TM_Uio (t-2184)	Tm_Gye (t-2184)	TM_Gye (t-2184)	load (t-2184)
day (t-2183)	Tm_Uio (t-2183)	TM_Uio (t-2183)	Tm_Gye (t-2183)	TM_Gye (t-2183)	load (t-2183)
...
day (t-1)	Tm_Uio (t-1)	TM_Uio (t-1)	Tm_Gye (t-1)	TM_Gye (t-1)	load (t-1)

Tabla 5.2. Series temporales utilizadas en el PDCP

5.1.2. Corrección de valores atípicos en el PDCP para días normales

Se utiliza el criterio de Chauvenet (ver ítem 2.2.4) para realizar el análisis de datos atípicos (feriados, fallas en el sistema de potencia, errores humanos, etc.).

Para el caso de la demanda no es posible aplicar el criterio directamente debido a las estacionalidades diaria y semanal de la serie temporal. La forma de análisis que se debe adoptar, considerando solo los datos de entrenamiento (13 semanas), consiste en separar la serie original en series conformadas por las observaciones correspondientes a cada periodo horario y para cada día de la semana, por lo que se obtendrán 7 matrices de días similares de 24 filas (horas) por 13 columnas (semanas). Se calculan también los valores medios horarios para cada día de la semana.

Los valores atípicos deben ser reemplazados para mejorar el desempeño del modelo, utilizando los siguientes criterios:

- Primeramente, se corrigen los días feriados por la media horaria correspondiente al promedio de los dos días laborales anteriores (si el feriado cae sobre día laboral).
- Si se tienen valores horarios consecutivos atípicos (factor inercial), se procede a reemplazar los valores horarios de todo el día por el promedio de los dos días similares anteriores. Se considera que esto ocurre con los días después de un feriado y no representan el comportamiento normal de la demanda.
- Cuando se trata de un solo valor atípico se reemplaza por el promedio entre los valores correspondientes a las horas anterior y posterior de dicho valor.
- Cuando se trate de valores atípicos con una duración mayor a 2 horas y menor a 4 horas (falla considerable en el sistema de potencia) se procede a reemplazar de la misma manera como se corrige el factor inercial.

5.1.3. Armado de los conjuntos de entrenamiento y validación o test

El error $E_{(t)}$ se calcula para la última semana de los datos de entrenamiento –semana de entrenamiento (168 periodos horarios)–, la cual está comprendida entre el 04 de septiembre y el 10 de septiembre de 2000. Para construir las matrices de comportamiento (Rule Base) a utilizar en el pronóstico de demanda de la primera hora del 04 de septiembre, se utilizan los datos horarios de demanda comprendidos entre el 12 de junio y el 03 de septiembre. Para el pronóstico de demanda de la segunda hora del 04 de septiembre, se utilizan los datos horarios de demanda comprendidos entre la segunda hora del 12 de junio y la primera hora del 04 de septiembre, y así sucesivamente hasta realizar el pronóstico de demanda de los 168 periodos horarios de entrenamiento

(1 hora de adelanto), según lo detallado en 4.2.4. En este proceso de entrenamiento (Aprendizaje Supervisado) se encuentra la máscara de correlación óptima que minimiza el error $E_{(t)}$ de la semana de entrenamiento y que se considera válida para el PDCP de la primera semana de test -del 11 al 17 de septiembre de 2000 (07 días de validación o test)-.

Para el PDCP del día lunes 11 de septiembre de 2000 (primer día de test), la base de datos que se utiliza para construir las matrices de comportamiento esta comprendida entre el 12 de junio y el 10 de septiembre de 2000 (13 semanas = 91 días).

Para el PDCP del día martes 12 de septiembre de 2000 (segundo día de test), la base de datos que se utiliza para construir las matrices de comportamiento esta comprendida entre el 12 de junio y el 11 de septiembre de 2000 (92 días). Sucesivamente se van incorporando los datos reales hasta que para el PDCP del último día de la semana de test (17 de septiembre), la base de datos que se utiliza para construir las matrices de comportamiento considera los datos históricos desde el 12 de junio hasta el 16 de septiembre de 2000.

5.1.4. Validez requerida de una máscara de correlación

Se considera que una misma máscara de correlación se utilizará para realizar el PDCP de las próximas 24 horas (1 día en adelanto) **durante toda una semana completa** (de Lunes a Domingo). Esto significa que:

- i) cada fin de semana (Domingo) debe ejecutarse el algoritmo de optimización para encontrar la máscara óptima de la semana entrante.
- ii) cada día (0 horas) debe ejecutarse el PDCP para las próximas 24 horas utilizando la máscara válida para esa semana y los datos reales registrados hasta el momento.

La validez semanal de la máscara de correlación será analizada en detalle en el ítem 5.2.7.

5.2. PDCP UTILIZANDO FIR & SRA

5.2.1. Análisis de la profundidad de la máscara de correlación

En la referencia [30], se utiliza como herramienta para encontrar la profundidad de la máscara la función de autocorrelación parcial (ACF). En la gráfica de la ACF se puede visualizar la correlación que existe en la serie temporal. En el caso que se tenga una correlación periódica (por ejemplo correlación semanal), la profundidad de la máscara debe analizar hasta por lo menos 3 periodicidades (profundidad máxima igual a 3 semanas).

Para el caso de esta investigación, la variable de interés que determinará la profundidad de la máscara será la demanda. En la Figura 2.6 se analiza la ACF de la demanda (05 de junio hasta el 10 de septiembre de 2000). Se grafican 338 retardos, que equivalen a 2

semanas. Si se analiza la gráfica, se determina que la demanda tiene 2 periodicidades: diaria y semanal.

Se tiene que tener en cuenta que con profundidades muy grandes, mayores que 168 (1 semana), el cálculo de optimización será muy complejo. Por ejemplo, si se asume que la máscara a analizar tiene una profundidad de 169 y se considera que existe 1 entrada para la variable tipo de día, 1 entrada para la temperatura y 8 entradas para la demanda, el número de posibles combinaciones será de $8.4487 E^{14}$, siendo este número casi 2 millones de veces mayor que el número de combinaciones considerando una profundidad de 25 (v_{10} de la Tabla 4.2). Por lo tanto no sería viable tener profundidades muy grandes para la máscara, por que esto resultaría en un modelo muy complejo de evaluar computacionalmente. Se concluye entonces que una búsqueda en este espacio de soluciones no sería factible.

Las profundidades de la máscara que en este trabajo se pretenden investigar son: 25, 49 y 73, es decir: uno, dos y tres días respectivamente. Como criterio general entonces, se analizará hasta una profundidad máxima equivalente a 3 periodicidades diarias para la demanda.

5.2.2. Número total de configuraciones a ser evaluadas por el SRA

Por lo ya expuesto, se asume que en la máscara óptima se tiene siempre 1 entrada en la variable tipo de día y 1 entrada para la variable temperatura máxima de la costa, y que no se consideran otras entradas de temperatura.

En la Tabla 5.3, se detalla el espacio de soluciones totales de la máscara de correlación planteada $|\Theta|$, variando la cantidad de entradas para la variable demanda entre 1 y 8, y considerando profundidades de máscaras de 25, 49 y 73.

ν	$ \Theta =X_A * X_E * \Theta _D$		
	profundidad=25	profundidad=49	profundidad=73
ν_3	15.000	115.248	383.688
ν_4	172.500	2.708.328	13.620.924
ν_5	1.265.000	41.527.696	317.821.560
ν_6	6.641.250	467.186.580	5.482.421.910
ν_7	26.565.000	4.111.241.904	74.560.937.976
ν_8	84.122.500	29.463.900.312	832.597.140.732
ν_9	216.315.000	176.783.401.872	7.850.201.612.616
ν_{10}	459.669.375	906.014.934.594	63.782.888.102.505
Total	73.663.741.824.070		

Tabla 5.3. Espacio de soluciones para una profundidad de 25, 49 y 73

5.2.3. Estructura del problema

Si analizamos la configuración ν_8 (dimensión 8 de la Tabla 4.2 correspondiente a 1 entrada para la variable tipo de día, 1 entrada para la variable temperatura máxima de la

costa y 6 entradas para la variable demanda), el problema de optimización debe formularse de la siguiente manera:

Minimizar:	$\text{Error_PDCP}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$
Tal que:	$1 \leq a_1, e_1 \leq 25$, $1 \leq f_1, f_2, f_3, f_4, f_5, f_6 \leq 24$ & $f_m \neq f_j$; $m=1, \dots, 6$ & $j \neq m$
donde;	a_1 : elemento de la variable tipo de día e_1 : elemento de la variable temperatura máxima de la Costa f_1, \dots, f_6 : elementos de la variable demanda

La función de evaluación o “función objetivo” (Error_PDCP), utiliza el error relativo para un periodo de una semana definido en la ecuación (4.4).

La metodología de solución que se plantea a este problema de optimización es utilizar el algoritmo evolutivo SRA, para encontrar una configuración óptima \mathbf{v}^* de la máscara de correlación, tal que minimice el error del PDCP para un grupo de datos de entrenamiento.

5.2.4. Hipótesis utilizadas en la simulación

Se deben simular todas las dimensiones (v_3, \dots, v_{10}) considerando que siempre existe una entrada de la variable “tipo de día” y que se incorpora solo una entrada para la variable “temperatura máxima de la Costa”, debido a la correlación significativa que ésta tiene con la demanda [43].

El valor del parámetro α (coeficiente de elasticidad) del algoritmo evolutivo SRA que se fija para el análisis en esta investigación es 0,92. Valor encontrado como óptimo en la referencia [24].

Otro parámetro que tiene que fijarse es el valor n de la estructura de vecindad propuesta. Se fija un valor de $n=6$ cuando la máscara tenga una profundidad de 25, un valor de $n=9$ cuando se tiene una profundidad de 49 y un valor de $n=12$ si la profundidad es de 73.

Con estas consideraciones se procede a simular las diferentes configuraciones planteadas para las semanas de test. En el apéndice D, se muestran los resultados obtenidos.

En la Tabla 5.4, se muestra un resumen de los resultados del apéndice D para las 2 semanas de test, con las 4 configuraciones que presentan el menor error promedio de entrenamiento $E_{(t)}$. (se muestran los errores de la primera semana de entrenamiento $-I_{tr}$ del 04 al 10 de septiembre de 2000 y de la segunda semana de entrenamiento $-II_{tr}$ del 11 al 17 de septiembre de 2000). Se observa también el error de cada periodo de test

$E_{TEST(t)}$ (2 semanas de validación, es decir del 11 de septiembre al 24 de septiembre de 2000).

ν	profundidad	$E(t)$		$E_{TEST(t)}$	
		I_{tr}	II_{tr}	I_{TEST}	II_{TEST}
6	25	1,2644	0,9655	1,2141	1,3909
7	25	1,2079	0,9129	1,0435	1,3447
8	25	1,1657	0,9122	0,9600	1,2760
9	25	1,1817	0,9211	0,9723	1,2981

Tabla 5.4. Resumen de los errores de entrenamiento ($E(t)$) y test ($E_{TEST(t)}$) obtenidos con las distintas configuraciones de ν^*

De todas las simulaciones realizadas, cuyos resultados se muestran en el apéndice D se concluye que la profundidad de la máscara óptima que conduce a menores errores de validación $E_{TEST(t)}$ es igual a 25, a excepción de la dimensión ν_4 , para la cual la profundidad óptima es igual a 49.

Analizando la Tabla 5.4, se observa que el menor error se obtiene con la dimensión 8 y es igual a $I_{tr}=1,1657\%$ / $I_{TEST}=0,9600\%$ y $II_{tr}=0,9122\%$ / $II_{TEST}=1,2760\%$.

5.2.5. Análisis del PDCP incluyendo otras Temperaturas

Hasta aquí se ha realizado el PDCP modelando solo la temperatura máxima de la Costa, debido a la correlación que tiene esta variable con la demanda [43]. En el apéndice E se muestran los resultados obtenidos **incluyendo 1 entrada para otra variable de temperatura**, que puede ser: temperatura mínima de la Sierra, temperatura máxima de la Sierra o temperatura mínima de la Costa, para las dimensiones $\nu = 7,8,9$ y 10; aumenta 1 dimensión a las dimensiones de la Tabla 5.4.

En la Tabla 5.5, se presentan los resultados y las configuraciones óptimas que alcanzaron el menor error de entrenamiento $E(t)$ para las diferentes temperaturas analizadas, correspondiendo todos los valores de la Tabla 5.5 a la inclusión de la temperatura mínima de la Costa (d_1). Se muestra también, para las dos semanas de validación, el error promedio del grupo de test $E_{TEST(t)}$.

ν	ν^*	$E(t)$		$E_{TEST(t)}$	
		I_{tr}	II_{tr}	I_{TEST}	II_{TEST}
7	$a_1, d_1, e_1, f_1, f_2, f_3, f_4$	1,1611	0,9571	1,2034	1,3478
8	$a_1, d_1, e_1, f_1, f_2, f_3, f_4, f_5$	1,1629	0,9480	1,2802	1,3625
9	$a_1, d_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6$	1,1928	0,9447	1,2696	1,4311
10	$a_1, d_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7$	1,1909	0,9483	1,1880	1,3711

Tabla 5.5. Resumen de los errores obtenidos con las distintas temperaturas

De la Tabla 5.5. se puede concluir que con la introducción en el modelo de una entrada más, que considere a la variable temperatura, no se disminuyen los errores del PDCP.

5.2.6. Análisis de sensibilidad de los k -NN

La metodología FIR propuesta por Cellier utiliza los cinco registros más semejantes (5-NN) para realizar el pronóstico fuzzy. En esta investigación se propone realizar un análisis de sensibilidad para el parámetro “ k -NN” variando k entre 4 y 6. En el apéndice F se muestran los resultados obtenidos para este análisis, considerando que la máscara de correlación tiene una dimensión de 5, 6, 7 y 8 (dimensiones con el menor error encontradas en el ítem 5.2.3).

La Tabla 5.6, presenta un resumen de los resultados del error de entrenamiento y de test para las configuraciones óptimas, con el k -NN correspondiente.

ν	k_{NN}	$E(t)$		$E_{\text{TEST}}(t)$	
		I_{tr}	II_{tr}	I_{TEST}	II_{TEST}
6	4	1,2348	0,8644	1,1970	1,2788
	6	1,3619	0,9926	1,2188	1,3498
7	4	1,1647	0,8737	1,0600	1,3557
	6	1,2663	0,9034	1,1025	1,3930
8	4	1,1615	0,8785	0,9400	1,2518
	6	1,1976	1,0778	0,9024	1,3432
9	4	1,1637	0,8856	0,9935	1,2767
	6	1,2118	0,9135	1,0165	1,3300

Tabla 5.6. Errores obtenidos con distintos k_{NN} para distintas dimensiones de ν

Se debe resaltar que el menor error de entrenamiento y test se obtiene con $k = 4$, para las distintas configuraciones simuladas, y que la configuración ν_8 produce los menores errores de entrenamiento y test. El mínimo error se obtiene por tanto para la configuración ν_8 y $k = 4$, con los errores siguientes $I_{\text{tr}}=1,1615\%$ / $I_{\text{TEST}}=0,9400\%$ y $II_{\text{tr}}=0,8785\%$ / $II_{\text{TEST}}=1,2518\%$.

En la Figura 5.1 se muestran los errores diarios relativos promedio para las dos semanas de test, incluyéndose los errores que se obtuvieron en la semana de entrenamiento respectiva, para la situación óptima (dimensión $\nu = 8$ y $k = 4$).

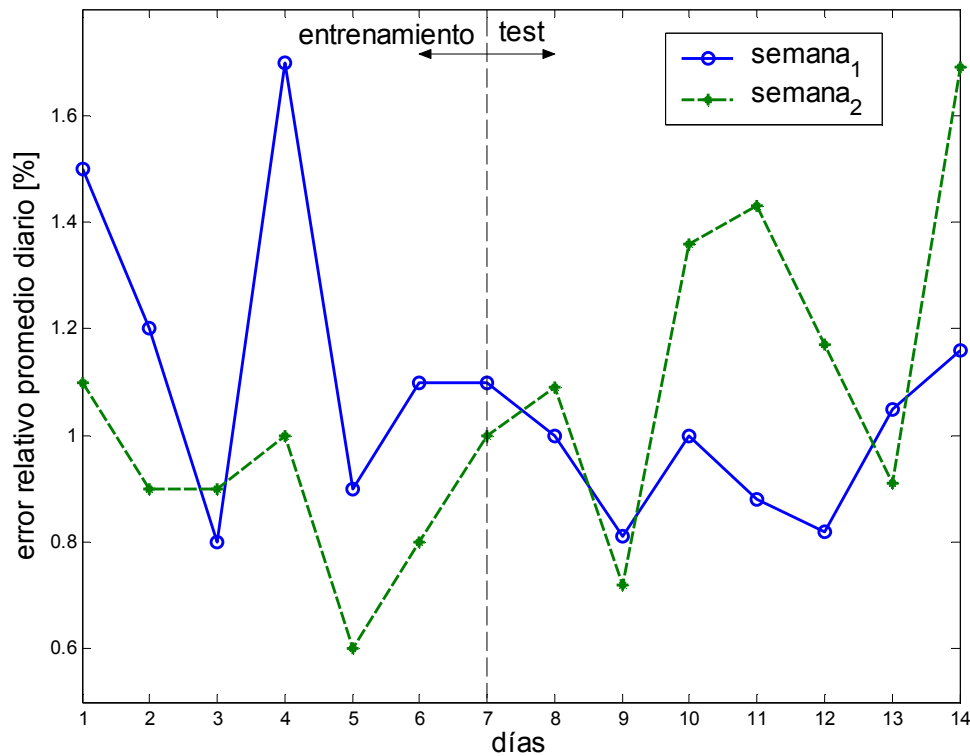


Figura 5.1. Errores diarios de entrenamiento y test para v_8 y 4_NN

5.2.7. Análisis de la validez de una máscara de correlación

En la Figura 5.1 se puede visualizar para las dos semanas de test (14 días), que los errores diarios de las predicciones poseen una desviación estándar acotada, es decir no existe una dispersión marcada a medida que se realiza el PDCP tanto para el día lunes como para el día domingo de cada una de las semanas.

En la Figura 5.2 se realiza el PDCP para 1 día en adelante considerando como semana de entrenamiento los datos de demanda comprendidos desde el 04 hasta el 10 de septiembre de 2000 y se obtiene el PDCP para 1 día en adelante para los 14 días de simulación (2 semanas de test comprendidas desde el 11 al 24 de septiembre de 2000), es decir se realiza la optimización de la función objetivo planteada en la ecuación (4.4) y con las entradas (máscara de correlación) producto de la optimización de dicha función objetivo se calcula el PDCP para los 14 días de test (horizonte de tiempo planteado como válido).

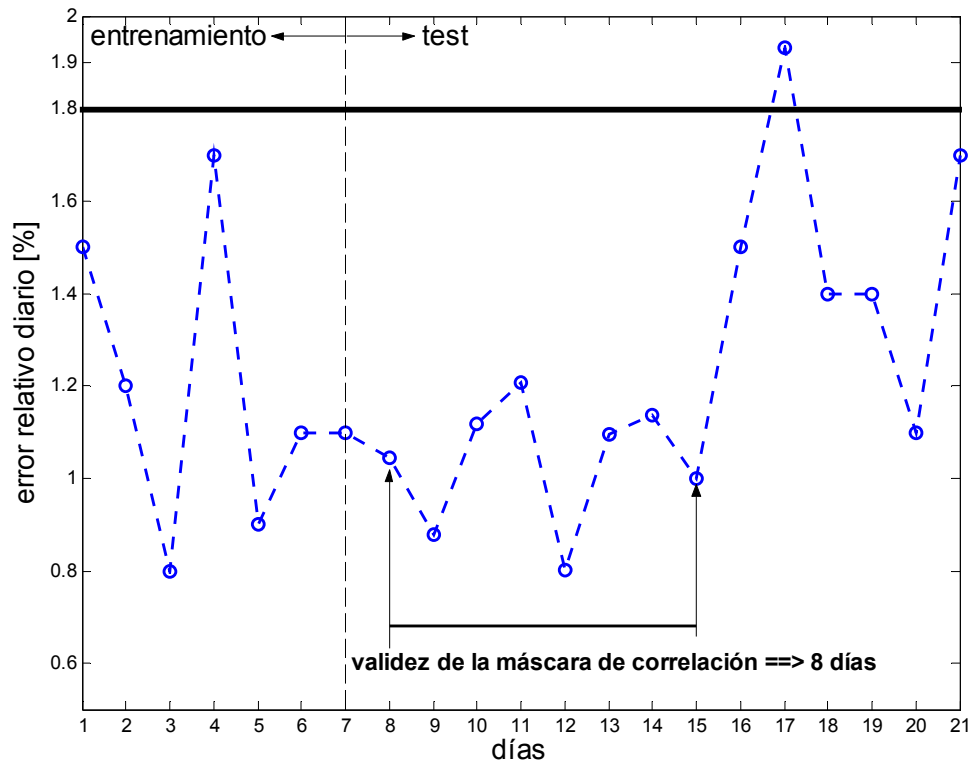


Figura 5.2. Errores diarios para el PDCP desde el 11 al 24 de septiembre de 2000

En la Figura 5.2 los errores relativos diarios de la semana de entrenamiento y de los siete primeros días de test (del 11 al 17 de septiembre de 2000) son los obtenidos anteriormente (Figura 5.1).

Puede observarse que la máscara de correlación es válida para los días de test desde el lunes 11 hasta el lunes 18 de septiembre de 2000. Los errores que se obtienen para la primera semana de test son aceptables, pero los errores que resultan del PDCP desde el martes 19 hasta el domingo 24 de septiembre de 2006 son altos y se observa que la tendencia es creciente a medida que el horizonte de pronóstico aumenta (aumenta la dispersión). De aquí que la máscara de correlación encontrada en el proceso de optimización (periodo de entrenamiento) tiene una validez de una semana. En otras palabras, si se realiza el PDCP para más de una semana utilizando la misma máscara de correlación, la desviación estándar no se mantiene acotada.

Se debe mencionar que la máscara de correlación encontrada en la optimización es diferente para cada uno de los periodos de entrenamiento. Este mismo problema ya se encontró en la referencia [6]. Una posible solución a este problema es documentado en la referencia [40].

5.3. ANÁLISIS DE PERFORMANCE DEL ALGORITMO SRA

Con el objeto de evaluar la sensibilidad del algoritmo de optimización utilizado para la determinación de la máscara óptima, se muestran los errores de entrenamiento (periodo del 04 al de 10 de septiembre de 2000) y test (periodo del 11 al de 17 de septiembre de

2000) obtenidos para la dimensión $v=8$ y para 4_NN, cuando se varía el parámetro α “coeficiente de elasticidad” del algoritmo SRA.

α	$E(t)$	$E_{TEST}(t)$	t [min]
0,72	0,9069	0,9400	96,68
0,75	0,9071	0,9984	101,32
0,78	0,9069	0,9400	117,58
0,81	0,9069	0,9400	95,70
0,84	0,9069	0,9400	90,28
0,87	0,9069	0,9400	110,70
0,90	0,9069	0,9400	103,30
0,93	0,9069	0,9400	66,57
0,96	0,9069	0,9400	55,50
0,99	0,9069	0,9400	39,12

Tabla 5.7. Análisis de sensibilidad para la primera semana de test del parámetro α

El tiempo óptimo para la simulación del algoritmo FIR & SRA es de 39,12 min y se obtiene con $\alpha = 0,99$. Con este coeficiente de elasticidad se encuentra el mínimo error de entrenamiento y test.

5.4. PDCP PARA OTRO PERIODO DE VALIDACIÓN

5.4.1. Base de Datos del Sistema Ejemplo

La segunda base de datos del sistema multivariable a validar en esta investigación, consiste en registros horarios de temperatura y demanda para el periodo desde el 13 de febrero hasta el 14 de mayo de 2006. Los grupos de datos de entrenamiento y de test se muestran en la Tabla 5.8.

Mes	lu	ma	mi	ju	vi	sa	do
febrero	13	14	15	16	17	18	19
	20	21	22	23	24	25	26
	27	28					
marzo			1	2	3	4	5
	6	7	8	9	10	11	12
	13	14	15	16	17	18	19
	20	21	22	23	24	25	26
abril	27	28	29	30	31		
						1	2
	3	4	5	6	7	8	9
	10	11	12	13	14	15	16
	17	18	19	20	21	22	23
mayo	24	25	26	27	28	29	30
	1	2	3	4	5	6	7
	8	9	10	11	12	13	14

Diagrama de anotación: Una gran llave a la izquierda agrupa los días del 13 de febrero al 14 de mayo como "Base de Datos de test II". Una llave a la derecha agrupa los días del 13 de febrero al 31 de marzo como "Base de Datos de entrenamiento". Flechas indican que los días 1-7 de mayo corresponden a la "Semana de entrenamiento" y los días 8-14 de mayo corresponden a la "Semana de test".

Tabla 5.8. Segunda Base de datos de entrenamiento y test

5.4.2. Armado de los conjuntos de entrenamiento y test

El error $E_{(t)}$ se calcula para la última semana de los datos de entrenamiento –semana de entrenamiento–, la cual está comprendida entre el 01 de mayo y el 07 de mayo de 2006. Para construir las matrices de comportamiento a utilizar en el pronóstico de demanda de la primera hora del 01 de mayo, se utilizan los datos horarios comprendidos entre el 13 de febrero y el 30 de abril. Para el pronóstico de la segunda hora del 01 de mayo, se utilizan los datos comprendidos entre la segunda hora del 13 de febrero y la primera hora del 01 de mayo, y así sucesivamente hasta realizar el pronóstico de demanda de los 168 periodos horarios, según lo detallado en 4.2.4. De este proceso de entrenamiento se obtiene la máscara óptima que minimiza el error $E_{(t)}$ para la semana de entrenamiento y que se considera válida para el PDCP del 08 al 14 de mayo de 2006 (semana de validación o test).

Para el PDCP del día lunes 08 de mayo de 2006, la base de datos que se utiliza para construir las matrices de comportamiento esta comprendida entre el 13 de febrero y el 07 de mayo de 2006. Para el PDCP del día martes 09 de mayo de 2006, la base de datos que se utiliza para construir las matrices de comportamiento esta comprendida entre el 13 de febrero y el 08 de mayo de 2006. Sucesivamente se van incorporando los datos reales hasta obtener el PDCP del último día de la semana de test (14 de mayo).

5.4.3. PDCP utilizando FIR & SRA

Hipótesis utilizadas en la simulación

- La profundidad óptima de la máscara de correlación es de 25 y el valor de $n=6$.
- La dimensión de la máscara de correlación es de v_8 (óptima en el proceso de simulación anterior). Esta dimensión considera una entrada de la variable “tipo de día”, una entrada para la variable “temperatura máxima de la Costa” y 6 entradas para la demanda.
- El valor del parámetro α del algoritmo evolutivo SRA es de 0,99.

Con estas consideraciones se procede a simular el modelo de pronóstico para la semana de test planteada. El menor error promedio de entrenamiento $E_{(t)}$. (del 01 al 07 de mayo de 2006) es igual a 1,0511% y el error de test $E_{TEST(t)}$ (del 08 al 14 de mayo de 2006) es igual a 1,3099%. Estos resultados confirman la precisión del nuevo método de pronóstico.

En la Tabla 5.9 se detallan la media y la desviación estándar del error de test, considerando las 3 semanas de test simuladas (del 11 al 24 de septiembre de 2000 y del 08 al 14 de mayo de 2006), resultantes de la aplicación de la Metodología FIR & SRA y su comparación con los resultados obtenidos con el método ARIMA (CENACE) y con las Redes Recurrentes de Jordan (RR). Se muestra además el error máximo obtenido en el PDCP con cada uno de los Métodos de Pronóstico antes mencionados.

Método de pronóstico	Media	Desviación	E max
ARIMA	1,5873	1,2617	8,0298
RR	1,4729	1,2064	7,1311
FIR & SRA	1,1672	1,0633	6,3374

Tabla 5.9. Media, desviación estándar y máximo error obtenido con distintos métodos de pronóstico

Se puede observar que el menor error de validación o test se obtiene con la Metodología FIR & SRA, reduciéndose el error relativo promedio del periodo de test en un 26,46% con respecto a la Metodología de Referencia ARIMA y un 20,75% con respecto a un Modelo de Redes Neuronales Artificiales (RR). Además con la Metodología FIR & SRA se reduce la desviación estándar del error de pronóstico (incertidumbre del resultado) en un 15,72%. Por lo tanto, la Metodología FIR & SRA obtiene menores errores de pronóstico y mayor certeza en el resultado.

En la Figura 5.3 se muestra la ACF de los errores de pronóstico para los periodos de test simulados. El Test de independencia de Anderson [4] acepta que el proceso es puramente aleatorio (ruido blanco) cuando la ACF se distribuye según una función normal $N(0,1/n)$, donde n es el tamaño de la muestra.

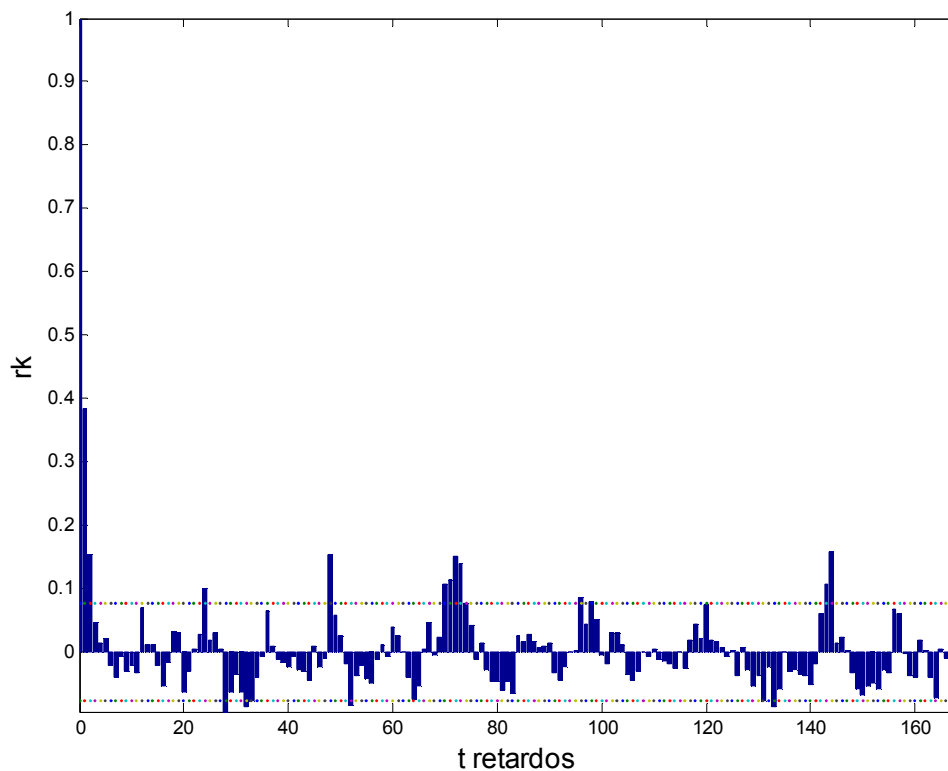


Figura 5.3. ACF de los residuos utilizando la Metodología FIR & SRA

En la gráfica se observa que para un límite de significación del 95% el error de pronóstico se aproxima a un proceso aleatorio.

5.5. EFECTOS ECONÓMICOS DEL PDCP EN EL DESPACHO ECONÓMICO

En esta sección se comparan los efectos económicos en la aplicación del Pronóstico de Demanda con horizonte semanal obtenido con la Metodología FIR & SRA, para el periodo de test del 08 al 14 de mayo de 2006, respecto del Pronóstico de Demanda semanal obtenido con la Metodología de Referencia (modelo ARIMA) que es utilizada en el CENACE.

Pronóstico de Demanda

En la Tabla 5.10. se visualizan los errores relativos promedios diarios utilizando la Metodología FIR & SRA y la Metodología de Referencia, considerando un horizonte de pronóstico de 168 periodos (1 semana). Se detalla además para la semana de test analizada, el porcentaje de disminución del error del Pronóstico de Demanda obtenido con la Metodología propuesta en esta Investigación.

Fecha	E FIR & SRA	E ARIMA
08/05/2006	1,2	2,2
09/05/2006	1,1	1,3
10/05/2006	1,7	1,2
11/05/2006	1,0	1,4
12/05/2006	1,5	1,3
13/05/2006	0,9	0,8
14/05/2006	1,6	2,8
promedio	1,3099	1,5873
reducción	17%	

Tabla 5.10. Errores relativos promedios del Pronóstico de Demanda

Los escenarios de demanda correspondientes a los errores relativos promedios de la Tabla 5.10. son los que se muestran en la Tabla 5.11.

Fecha	Energía [MWh]		
	Real	FIR & SRA	ARIMA
08/05/2006	41.421	41.175	41.739
09/05/2006	42.934	42.892	43.305
10/05/2006	43.395	43.981	43.370
11/05/2006	43.194	42.966	42.954
12/05/2006	43.132	42.531	42.792
13/05/2006	38.701	38.805	38.659
14/05/2006	34.913	34.393	35.907
total	287.689	286.744	288.727

Tabla 5.11. Escenarios de demanda considerados en el análisis

En la Figura 5.4. se muestra para el día lunes 08 de mayo de 2006, la demanda real, la demanda pronosticada con la Metodología ARIMA y la demanda pronosticada con la Metodología FIR & SRA. Se adjunta además, para cada Metodología, la diferencia entre el pronóstico de demanda y la demanda real. Esta diferencia nos da una idea clara de los errores que se cometieron en el pronóstico, tanto para las horas que se produce sobreestimación como para las horas en subestimación.

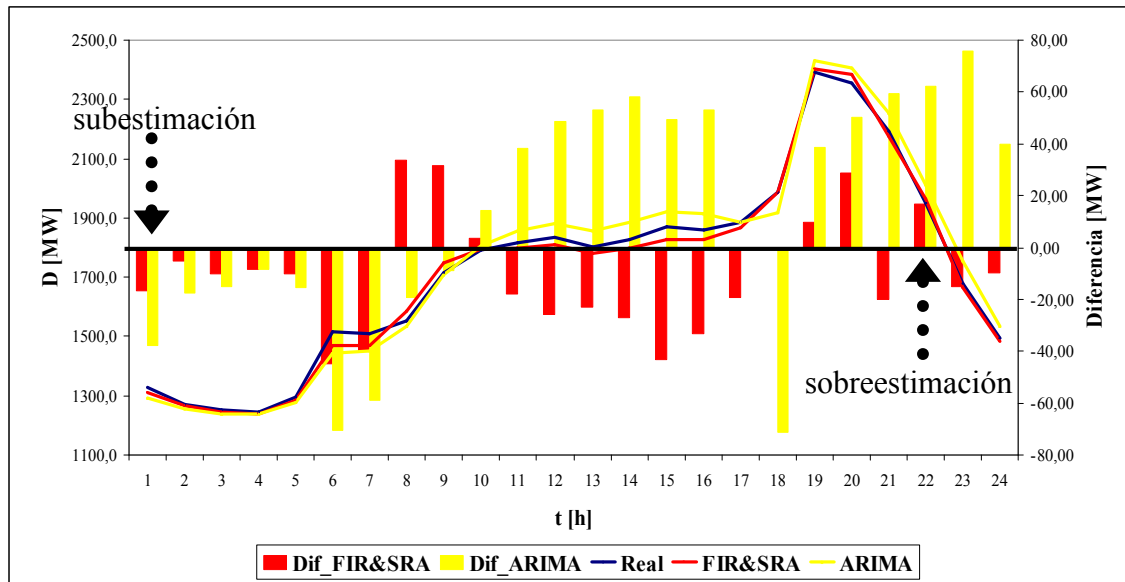


Figura 5.4. Análisis del pronóstico de demanda y de sus desvíos

Como ejemplo, se analiza lo que ocurre el día lunes 08 de mayo de 2006 (ver Figura 5.4.):

- $t=22:00$ (los dos modelos sobreestiman la demanda). La demanda real del SSEE del Ecuador fue de 1949,4 MW, el modelo FIR & SRA pronosticó la demanda igual a 1966,3 MW y el modelo ARIMA obtuvo un pronóstico de demanda de 2011,7 MW. El primer modelo analizado sobrestimó la demanda en 16,9 MW, mientras que el modelo utilizado en el CENACE sobrestimó la demanda en 62,3 MW.
- $t=01:00$ (los dos modelos subestiman la demanda). La demanda real del SSEE del Ecuador fue de 1328,0 MW, el modelo FIR & SRA pronosticó la demanda igual a 1311,1 MW y el modelo ARIMA obtuvo un pronóstico de demanda de 1290,3 MW. El primer modelo subestima la demanda en 16,9 MW, mientras que el modelo utilizado en el CENACE subestimó la demanda en 37,7 MW.

Para la semana de test analizada se obtuvo con la Metodología FIR & SRA una sobreestimación de energía de 1444 MWh y una subestimación de 1776 MWh. Para la misma semana, la Metodología ARIMA tuvo una sobreestimación de energía de 2813 MWh y una subestimación de 2389 MWh.

A continuación se analizará el efecto económico de las desviaciones en el pronóstico de demanda (subestimación y sobrestimación de potencia/energía) en el Despacho Económico del Sistema Eléctrico Ecuatoriano.

Despacho Económico

Para el análisis de los efectos económicos del Pronóstico de Demanda en la Programación Semanal, se debe simular para los distintos escenarios de demanda la programación del parque generador.

Para la Programación Semanal se utiliza el Modelo de Despacho “Jasson³” utilizado por la Dirección de Planeamiento (DPL) del CENACE. Con este programa se obtiene el despacho semanal del parque hidrotérmico del SSEE del Ecuador.

Características del Modelo de Despacho

- El objetivo es minimizar el costo de producción en barra de mercado
- El modelo es uninodal. Esta única barra representa el centro de carga del Sistema Nacional Interconectado donde ocurre la demanda y oferta de energía y potencia.
- La demanda se considera de naturaleza determinística y se representa a potencia constante en bloques horarios en un período de 168 horas.
- Las unidades térmicas se consideran separadamente y se simulan dentro de sus parámetros técnicos de operación. La indisponibilidad (mantenimientos programados o correctivos) se consideran dentro de los períodos horarios.
- Se representa la generación de las centrales de pasada.
- Las centrales hidroeléctricas con capacidad de almacenamiento se simulan dentro de sus parámetros técnicos de operación más la asignación de una cuota diaria energética que representa la disponibilidad de energía para cada día asignada a cada una de éstas centrales y cuyo valor viene establecido por la política de operación de largo plazo desagregada en períodos semanales.
- Se toman en cuenta las restricciones de la red de transmisión así como sus mantenimientos programados y/o correctivos en períodos horarios.
- Se toma en cuenta la importación o exportación de energía a través de las interconexiones internacionales.

Características de la oferta del SSEE ecuatoriano

³ Módulo computacional utilizado en el CENACE y que sirve de interfaz para la ejecución de la Programación Semanal. Este módulo consiste en la ejecución de despachos diarios mediante un modelo basado en un algoritmo de utilización en redes.

El parque hidroeléctrico y los costos variables de producción del parque termoeléctrico para el mes de mayo de 2006 se muestran en el apéndice G.

Procederemos a analizar el parque generador real del sistema ecuatoriano. En la Figura 5.5. se muestran para el mes de mayo de 2006, las potencias efectivas del parque hidrotérmico para cada una de las tecnologías presentes.

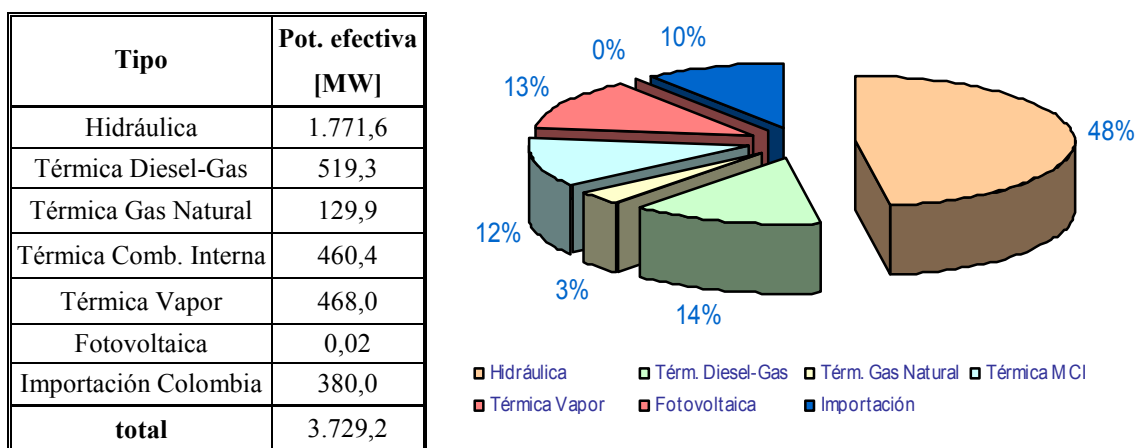


Figura 5.5. Características de la oferta del SSEE ecuatoriano

Para la semana de test planteada, desde el 08 al 14 de mayo de 2006, la producción hidráulica real fue de 160,281 GWh que representa el 56% de la energía real requerida para abastecer la demanda del sistema ecuatoriano (287,689 GWh).

Para dicha semana la producción de la central hidroeléctrica Paute (central de embalse) fue de 126,832 GWh, que representa el 44% de la demanda ecuatoriana. En el SSEE del Ecuador la central Paute es la encargada de cubrir las desviaciones que se producen debido al pronóstico de demanda, es decir es la unidad asignada para Regulación Secundaria de Frecuencia. El embalse de esta central tiene una regulación semanal, como consecuencia su operación puede afectar la evolución de los precios diarios del Mercado.

Análisis Económico

Para analizar los efectos económicos del Pronóstico de Demanda se realiza la simulación del despacho económico para los escenarios de demanda propuestos, considerando las siguientes hipótesis:

1. Se considerará que el costo de producción real del sistema se produce cuando en el despacho semanal se simula la demanda real.

Ahora, lo que se requiere analizar son los efectos económicos en el despacho al realizar la optimización con un Pronóstico de Demanda (Metodología FIR & SRA y

Metodología ARIMA), pronóstico que tiene inmerso un error producto de las incertidumbres presentes.

2. Se considera fija la energía hidráulica de pasada y embalse, es decir para cada escenario se asume igual cota energética y el mismo despacho horario (producción hidráulica obtenida con la simulación de la demanda real en la Programación Semanal). Entonces, el análisis se basa principalmente en la asignación de unidades térmicas.
3. Cuando el modelo de pronóstico de demanda produce sobreestimación debido al error de pronóstico, se crea un exceso en la asignación de generación en el SSEE debido a que la demanda real es menor, produciéndose un sobrecosto debido al ingreso innecesario de unidades caras que no son requeridas.
4. Cuando el modelo de pronóstico de demanda produce en cambio subestimación se crea un déficit de energía, ya que la demanda real es mayor, por lo que para poder abastecer esta diferencia de demanda se tiene que realizar un “Redespacho por Reserva”, que por ejemplo considera el ingreso de unidades rápidas – del orden de los minutos – que se encuentran en reserva fría en el SSEE, como las turbogas.

Para poder evaluar el sobrecosto producido por este redespacho se utilizará un costo variable de producción de una unidad turbogas de 90 US\$/MWh (costo variable declarado para la unidad 6 de la empresa CATEG – generación).

Finalmente con estas hipótesis, se debe analizar las consecuencias de ambos despachos (FIR & SRA y ARIMA), confrontados con el despacho real del sistema.

Para poder mostrar los efectos económicos del pronóstico de demanda en el despacho económico, se realiza a modo de ejemplo un análisis para dos horas del día lunes 08 de mayo de 2006:

- $t=22:00$ (los dos modelos sobreestiman la demanda).

El despacho de referencia y el costo de producción real para el SSEE fue de US\$ 24.942,9 dólares (ver Tabla 5.12.). La unidad marginal es la TV2 de la central Gonzalo Zevallos (57,125 US\$/MWh).

central	Costo variable [US\$/MWh]	Despacho [MW]			Predespacho D. real		Costo [US\$]		
		REAL	FIR & SRA	ARIMA	FIR & SRA	ARIMA	REAL	FIR&SRA	ARIMA
Hidráulico	2 ⁴	1420,0	1420,0	1420,0	1420,0	1420,0	2840,0	2840,0	2840,0
Esmeraldas	27,038	132,5	132,5	132,5	132,5	132,5	3582,5	3582,5	3582,5
Guangopolo 1	41,748	5,2	5,2	5,2	5,2	5,2	217,1	217,1	217,1
Guangopolo 2	42,674	5,2	5,2	5,2	5,2	5,2	221,9	221,9	221,9
Guangopolo 5	42,728	5,2	5,2	5,2	5,2	5,2	222,2	222,2	222,2
Guangopolo 6	42,783	5,2	5,2	5,2	5,2	5,2	222,5	222,5	222,5
Guangopolo 3	42,837	5,2	5,2	5,2	5,2	5,2	222,8	222,8	222,8
Guangopolo 4	42,892	5,2	5,2	5,2	5,2	5,2	223,0	223,0	223,0
Machala Power A	43,887	65,0	65,0	65,0	65,0	65,0	2852,7	2852,7	2852,7
Machala Power B	44,873	64,9	64,9	64,9	64,9	64,9	2912,3	2912,3	2912,3
El Descanso 3	45,902	4,3	4,3	4,3	4,3	4,3	197,4	197,4	197,4
G. Hernández 2	45,981	5,2	5,2	5,2	5,2	5,2	239,1	239,1	239,1
G. Hernández 1	46,273	5,2	5,2	5,2	5,2	5,2	240,6	240,6	240,6
G. Hernández 6	46,273	5,2	5,2	5,2	5,2	5,2	240,6	240,6	240,6
G. Hernández 4	46,379	5,2	5,2	5,2	5,2	5,2	241,2	241,2	241,2
G. Hernández 5	46,464	5,2	5,2	5,2	5,2	5,2	241,6	241,6	241,6
G. Hernández 3	46,829	4,3	4,3	4,3	4,3	4,3	201,4	201,4	201,4
El Descanso 4	45,910	5,2	5,2	5,2	5,2	5,2	238,7	238,7	238,7
C. S. Alegre	47,163	5	5	5	5	5	235,8	235,8	235,8
Trinitaria	45,983	133	133	133	133	133	6115,7	6115,7	6115,7
El Descanso 2	47,753	4,3	4,3	4,3	4,3	4,3	205,3	205,3	205,3
El Descanso 1	48,032	4,3	4,3	4,3	4,3	4,3	206,5	206,5	206,5
G. Zevallos TV2	57,125	49,4	66,3	71,2	49,4	31,4	2822,0	2822,0	1793,7
G. Zevallos TV3 ⁵	57,411			40,5		18,0			1033,4
TOTAL		1949,4	1966,3	2011,7	1949,4	1949,4	24942,9	24942,9	24948,0

Tabla 5.12. Análisis económico para sobreestimación considerando el despacho y predespacho

De las columnas (4) y (5) se determina el predespacho que se obtiene considerando el modelo de pronóstico de demanda FIR & SRA y ARIMA respectivamente. Ahora, si se compara los costos operativos en que se incurre con ambos predespachos, cuando se debe cubrir la demanda real, resulta que con la Metodología FIR & SRA no se produce ningún sobrecosto por el error de pronóstico, ya que considera en el despacho a la misma unidad (pero con mayor generación) que en el despacho de referencia. En cambio con la Metodología ARIMA se obtiene un sobrecosto de (24.948,0 menos 24.942,9) US\$ 5,1 dólares, ya

⁴ El costo variable de producción para una central hidráulica está dado por el costo medio representativo de operación y mantenimiento (2 US\$/MWh).

⁵ Las unidades TV2 y TV3 de la central Gonzalo Zevallos tienen una restricción técnica para la potencia mínima de 18 MW.

que la unidad TV3 queda despachada a mínimo técnico, absorbiendo parte de la generación que en el despacho real tenía la unidad TV2 de la central Gonzalo Zevallos.

A esta hora el CENACE si hubiera utilizado el modelo de pronóstico FIR & SRA para estimar la demanda, obtendría un ahorro de **US\$ 5,1 dólares**. Por lo que, el ahorro referido al costo real que se obtendría en el SSEE ecuatoriano si se utiliza la Metodología FIR & SRA en esta hora sería de (5,1/24.942,9) 0,02%.

- t=01:00 (los dos modelos subestiman la demanda).

El despacho real y los despachos para la Metodología FIR & SRA y ARIMA se muestran en la Tabla 5.13. La unidad marginal es la central Trinitaria (45,983 US\$/MWh).

central	Costo variable [US\$/MWh]	Despacho [MW]		
		REAL	FIR & SRA	ARIMA
Hidráulico	2	893,9	893,9	893,9
Esmeraldas	27,038	132,5	132,5	132,5
Guangopolo 1	41,748	5,2	5,2	5,2
Guangopolo 2	42,674	5,2	5,2	5,2
Guangopolo 5	42,728	5,2	5,2	5,2
Guangopolo 6	42,783	5,2	5,2	5,2
Guangopolo 3	42,837	5,2	5,2	5,2
Guangopolo 4	42,892	5,2	5,2	5,2
Machala Power A	43,887	65,0	65,0	65,0
Machala Power B	44,873	64,9	64,9	64,9
El Descanso 3	45,902	4,3	4,3	4,3
G. Hernández 2	45,981	5,2	5,2	5,2
G. Hernández 1	46,273	5,2	5,2	5,2
G. Hernández 6	46,273	5,2	5,2	5,2
G. Hernández 4	46,379	5,2	5,2	5,2
G. Hernández 5	46,464	5,2	5,2	5,2
G. Hernández 3	46,829	4,3	4,3	4,3
El Descanso 4	45,910	5,2	5,2	5,2
C. S. Alegre	47,163	5	5	5
Trinitaria	45,983	95,7	78,8	58
TOTAL		1328,0	1311,1	1290,3

Tabla 5.13. Análisis económico para subestimación considerando el despacho y predespacho

El costo de producción real para el SSEE fue de US\$ 18.941,7 dólares. Si se procede a utilizar el predespacho obtenido con las Metodologías FIR & SRA y ARIMA, resulta que la subestimación del modelo de pronóstico ocasiona un déficit

de generación, por lo que el operador del sistema debe realizar un redespacho por reserva para poder abastecer la demanda real del SSEE o en el peor de los casos seccionar carga con el respectivo costo por energía no suministrada.

El modelo FIR & SRA subestimó la demanda en 16,9 MW. Para abastecer esta demanda se debe realizar un redespacho por reserva, cuyo costo es de $(16,9 \times 90,0)$ US\$ 1521,0 dólares y para estimar el sobrecosto se debe considerar además, que la unidad marginal podría generar el correspondiente déficit a un costo de $(16,9 \times 45,983)$ US\$ 771,1 dólares, por lo tanto el sobrecosto producido por este modelo es de $(1521,0 - 771,1)$ US\$ 743,9 dólares.

El modelo utilizado en el CENACE subestimó la demanda en 37,7 MW. Para abastecer esta demanda se debe realizar un redespacho por reserva, cuyo costo es de $(37,7 \times 90,0)$ US\$ 3393,0 dólares. La central Trinitaria tiene una potencia efectiva de 133 MW, por lo que no podría abastecer el déficit de demanda $(95,7 + 37,7 = 133,4)$, entonces los 0,4 MW sería generados por la unidad 2 de la central El Descanso (costo variable de 47,753 US\$/MW), a un costo de $(37,3 \times 45,983 + 0,4 \times 47,753)$ US\$ 1734,3 dólares, por lo tanto el sobrecosto producido por este modelo es de $(3393,0 - 1734,3)$ US\$ 1.658,7 dólares.

A esta hora el CENACE si hubiera utilizado el modelo de pronóstico FIR & SRA para estimar la demanda, obtendría un ahorro de **US\$ 914,8 dólares**. Por lo que, el ahorro referido al costo real que se obtendría en el SSEE ecuatoriano si se utiliza la Metodología FIR & SRA en esta hora sería de $(914,8 / 18.941,7)$ 4,83%.

Con la Metodología FIR & SRA se obtuvo, para la semana de test analizada, una sobreestimación de 1444,0 MWh que ocasiona un sobrecosto en la operación real de US\$ 1.860 dólares, mientras que para subestimación se obtuvo una energía de 1776,0 MWh que ocasiona un sobrecosto de US\$ 52.706 dólares.

La Metodología ARIMA, tuvo una sobreestimación de 2813,0 MWh que ocasiona un sobrecosto en la operación de US\$ 3.951 dólares y para subestimación el modelo tuvo una energía de 2389,0 MWh que ocasiona un sobrecosto de US\$ 77.686 dólares.

Entonces, si en la Dirección de Planeamiento del CENACE se hubiera realizado el Pronóstico de Demanda semanal con la Metodología FIR & SRA, se obtendría un ahorro para el Mercado Eléctrico del Ecuador de US\$ 27.071 dólares.

En la Tabla 5.14. , se muestra el costo de producción cuando se ha simulado el despacho semanal para la demanda real.

fecha	Costo de producción [US\$]	sobrecosto FIR & SRA [US\$]				sobrecosto ARIMA [US\$]			
		energía	reserva	total		energía	reserva	total	
08/05/2006	671.394	57	11.876	11.933	1,78%	89	12.574	12.663	1,89%
09/05/2006	912.328	56	2.097	2.153	0,24%	150	6.004	6.154	0,67%
10/05/2006	927.366	100	6.342	6.442	0,69%	61	1.659	1.720	0,19%
11/05/2006	720.176	1.501	8.962	10.463	1,45%	2.258	12.616	14.874	2,07%
12/05/2006	721.324	5	15.012	15.017	2,08%	6	17.639	17.645	2,45%
13/05/2006	604.075	2	4.527	4.529	0,75%	3	6.055	6.058	1,00%
14/05/2006	439.307	140	3.890	4.030	0,92%	1.383	21.139	22.522	5,13%
total	4.995.970	1.860	52.706	54.566	1,09%	3.951	77.686	81.637	1,63%

Tabla 5.14. Análisis de los sobrecostos para las Metodologías FIR & SRA y ARIMA

Por lo tanto, el ahorro que se obtendría para la semana de test planteada en el SSEE ecuatoriano si se utiliza la Metodología FIR & SRA, sería de 0,54% con respecto al costo de producción real.

5.6. APLICACIÓN DE LA METODOLOGÍA FIR A LA PLANIFICACIÓN DE LA OPERACIÓN EN EL SSEE DEL ECUADOR (CENACE)

Como se puede observar en el apéndice H el pronóstico de demanda es una tarea de fundamental importancia para el Operador del Mercado.

La Dirección de Planeamiento del CENACE debe realizar la Programación de la Operación de Mediano y Corto plazo (de 24 horas hasta 168 horas en adelante), para lo cual requiere:

- Pronóstico de demanda con 1 semana de adelanto es decir necesita el pronóstico de demanda con 168 horas de adelanto (mediano plazo): para la Programación Semanal.
- Pronóstico de demanda con 24 horas de adelanto: para la Programación Diaria.

La Dirección de Operación del CENACE debe realizar la Programación de la Operación de Muy Corto Plazo (de 1 a 24 horas de adelanto), para lo cual requiere:

- Pronóstico de demanda desde 1 a varias horas en adelante.

5.6.1. Esquema del Procedimiento para el Pronóstico de Demanda

Para este análisis se considera como periodo de test, los datos horarios de demanda y temperatura del periodo comprendido desde el 11 al 17 de septiembre de 2000. El procedimiento a implementar es el siguiente:

- i) Se realiza la optimización de la función objetivo utilizando la ecuación (4.4), utilizando como semana de entrenamiento los datos de demanda desde el 04 hasta el 10 de septiembre de 2000.

- ii) Se obtiene el pronóstico de demanda con horizonte de 168 horas en adelante (pronóstico semanal).

Con la máscara de correlación óptima (del ítem anterior) se realiza el pronóstico semanal desde la 01:00 del 11 de septiembre hasta las 24:00 del 17 de septiembre de 2000. Dicha máscara -como se ha demostrado en esta Investigación- es válida para pronosticar un horizonte semanal. Este pronóstico se utiliza como entrada para la Programación Semanal.

- iii) Para la Programación Diaria, se obtiene el pronóstico de demanda con horizonte de 24 horas en adelante (pronóstico diario).

El pronóstico de demanda para el día lunes 11 de septiembre va a coincidir con el pronóstico semanal del día lunes, la diferencia radica en que una vez que se termina el día objetivo de planificación (por ejemplo el lunes 11), los datos reales de demanda de dicho día se utilizan para pronosticar el día siguiente (martes 12) y así sucesivamente hasta cumplir los 7 días de test en que la máscara de correlación es válida.

- iv) Esta herramienta de análisis servirá como metodología de pronóstico de demanda de muy corto plazo para la Supervisión y Control de la Operación con horizontes de 1 hasta varias horas en adelante.

Con este análisis se puede, incorporando datos reales de las últimas horas, observar una actualización del comportamiento para las siguientes horas del día (mejor estimación de la demanda real) y si fuera del caso -cuando se producen desviaciones considerables de demanda- realizar un Redespacho por Desviaciones de Demanda y así optimizar el costo operativo y la Reserva del Sistema.

En la Figura 5.6 se muestra la demanda real para la semana del 11 al 15 de septiembre de 2000. Además se incluye: el pronóstico semanal de la demanda y el pronóstico diario de la demanda de dicho periodo utilizando la Metodología FIR & SRA y el procedimiento descrito más arriba.

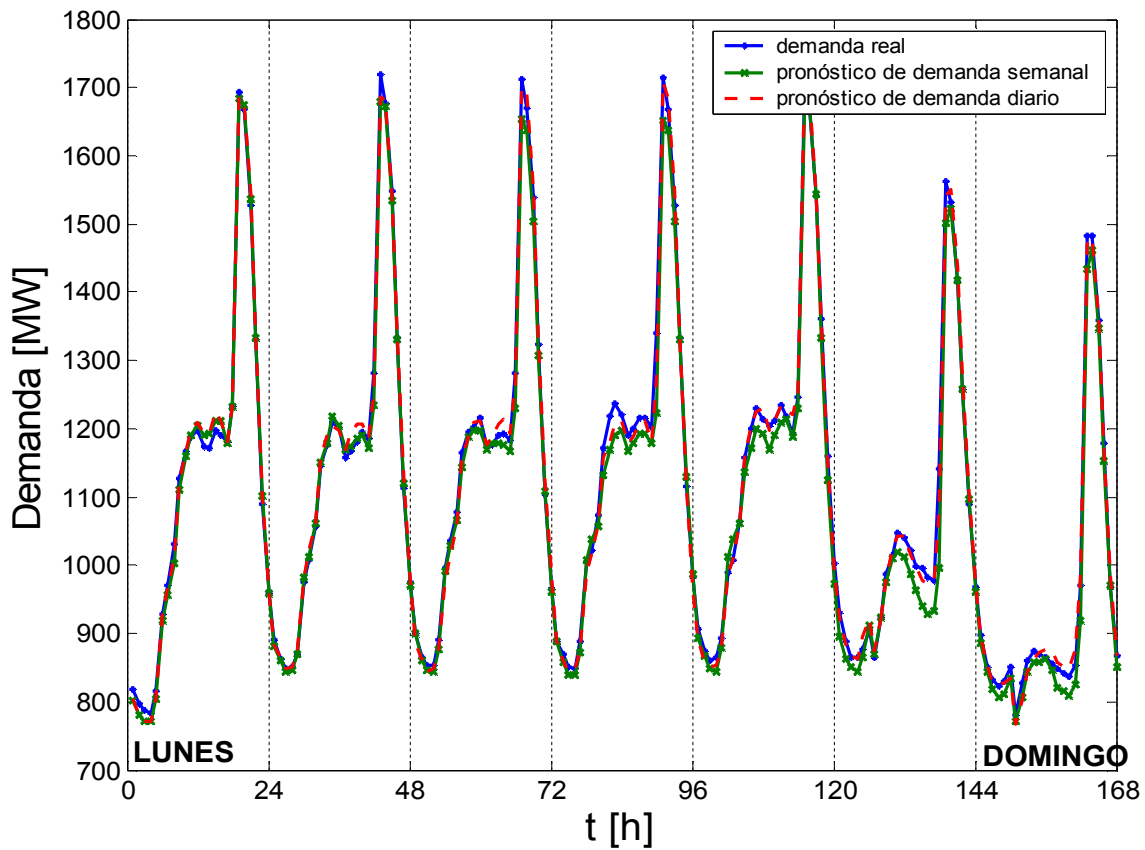


Figura 5.6. Pronóstico Semanal y Diario para la semana del 11 al 17 de septiembre de 2000

En la Tabla 5.15, se muestran los errores relativos diarios obtenidos para la semana de test, considerando el pronóstico semanal y el pronóstico diario.

Pronóstico	lunes 11	martes 12	miércoles 13	jueves 14	viernes 15	sábado 16	domingo 17	promedio
Semanal	1,0648	0,7198	1,1964	2,0641	1,7131	2,5574	2,0359	1,6217
Diario	1,0648	0,6369	1,0537	1,1936	0,7028	1,0193	0,9082	0,9400

Tabla 5.15. Error relativo diario considerando el pronóstico de demanda semanal y diario

Se puede observar claramente que en el pronóstico semanal los errores aumentan a medida que se avanza en el pronóstico, debido a la incertidumbre asociada a un mayor horizonte de observación. Estas desviaciones afectarán la Programación Semanal que se realiza al inicio de la semana y se producen debido a factores no considerados en la estimación de la demanda semanal, las cuales con un alcance diario se corrigen a medida que sucede el día de optimización en la Programación Diaria (incorporación a la Base de Conocimiento de los datos del último día, para obtener el pronóstico de demanda del siguiente día).

En la Tabla 5.15 se observa que el 14 de septiembre es el día de mayor desvío (1,1936%). Se realiza un análisis para el pronóstico de demanda de muy corto plazo

(PDMCP con horizontes de 15, 10 y 6 horas en adelante) tomando como referencia en la demanda media diferentes horas de inicio del pronóstico.

En la Figura 5.7 se muestran la demanda real del día jueves 14 de septiembre de 2000 y el PDMCP (pronóstico de demanda de muy corto plazo) con horizontes de: 15, 10 y 6 horas en adelante.

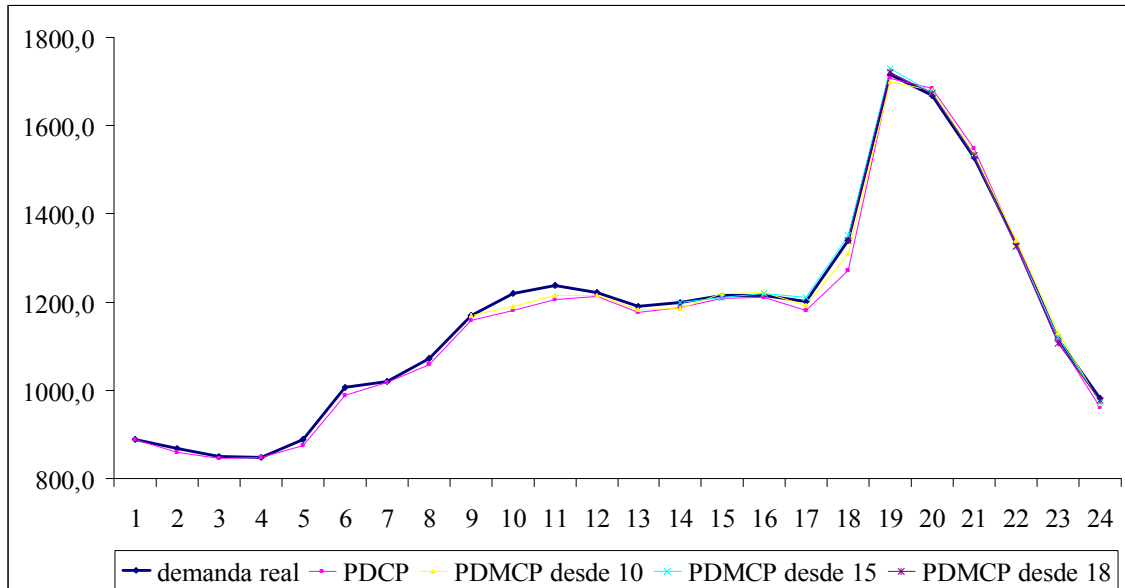


Figura 5.7. Pronóstico diario y horario para el día jueves 14 de septiembre de 2000

En la Tabla 5.16, se muestran los errores relativos horarios para el jueves 14 de septiembre de 2000, considerando el pronóstico diario y varios horizontes de tiempo para el pronóstico horario. Tanto en la Figura 5.7 como en la Tabla 5.16 se puede observar claramente como el error de pronóstico disminuye a medida que se incorporan nuevos datos a la Base de Datos de referencia.

jueves 14	REAL	FIR_D	FIR_10	FIR_15	FIR_18	PDCP	e_FIR_10	e_FIR_15	e_FIR_18
01:00	888,9	888,9				0,00			
02:00	868,8	859,3				1,09			
03:00	850,9	844,4				0,76			
04:00	848,3	847,8				0,06			
05:00	887,4	873,9				1,52			
06:00	1005,4	988,9				1,64			
07:00	1020,7	1018,1				0,25			
08:00	1072,7	1058,6				1,31			
09:00	1170,5	1158,2	1170,5			1,05			
10:00	1219,3	1180,2	1190,2			3,21	2,39		
11:00	1236,7	1206,0	1216,0			2,48	1,67		
12:00	1221,3	1212,4	1215,0			0,73	0,52		
13:00	1190,8	1177,5	1183,0			1,12	0,66		
14:00	1199,7	1187,8	1185,0	1199,7		0,99	1,23		
15:00	1215,6	1209,0	1219,0	1210,0		0,54	0,28	0,46	
16:00	1214,9	1211,0	1221,0	1220,0		0,32	0,50	0,42	
17:00	1202,4	1181,5	1191,5	1210,0		1,74	0,91	0,63	
18:00	1339,4	1272,4	1310,0	1350,0	1339,4	4,68	2,20	0,79	
19:00	1715,1	1706,7	1700,0	1730,0	1720,0	0,49	0,88	0,87	0,29
20:00	1668,4	1684,9	1679,9	1674,9	1673,0	0,99	0,69	0,39	0,28
21:00	1527,8	1547,7	1537,7	1532,0	1532,0	1,30	0,65	0,27	0,27
22:00	1332,5	1334,2	1340,0	1325,0	1325,0	0,13	0,56	0,56	0,56
23:00	1114,6	1113,6	1130,0	1120,0	1105,0	0,09	1,38	0,48	0,86
00:00	982,2	961,2	971,2	973,0	977,0	2,14	1,12	0,94	0,53
e promedio						1,1936	1,043	0,581	0,465

Tabla 5.16. Error relativo considerando varios horizontes horarios para el PDMCP

También se puede observar que la demanda pronosticada con horizonte diario (PDCP) tiene desviaciones (subestimación de la demanda) debido por ejemplo al incremento de la temperatura en la ciudad de Guayaquil, factor que no fue considerado el día miércoles 13 de septiembre cuando se realizó el PDCP para el jueves 14 de septiembre de 2000.

Estas desviaciones afectan la Programación Diaria, absorbiendo dichas desviaciones la central hidroeléctrica Paute (unidad asignada para realizar Regulación Secundaria de Frecuencia – RSF) cuando se realiza la Operación en Tiempo Real. Se debe analizar también si en el pico no se verá afectada la reserva del sistema.

Esta modelación será utilizada por la Dirección de Operaciones, la cual es encargada de la Supervisión y Control del Sistema Eléctrico Nacional.

La ventaja del método propuesto es la incorporación de un patrón más actualizado del comportamiento real de la demanda, disminuyendo el error de pronóstico a medida que se incorporan más datos reales de demanda (diferentes horas de inicio del PDMCP), como se puede ver con el error promedio obtenido para los tres horizontes de tiempo definidos (Tabla 5.16).

Otra aplicación práctica, será prever de una manera más acertada el comportamiento de la demanda pico. Por ejemplo se requiere un pronóstico desde las 19:00 en adelante (se considera que la demanda pico se produce en el periodo desde las 19:00 hasta las 22:00).

Con la incorporación de los últimos datos de demanda real (hasta las 18:00), ayudamos al modelo a la predicción de la demanda pico, por lo que se podrá analizar como se verá afectado el sistema cuando se produzca la máxima demanda y si es del caso prever el arranque de unidades por Reserva.

6. RESUMEN FINAL Y CONCLUSIONES

El pronóstico de demanda de corto plazo hoy en día es una de las tareas más importantes en lo que respecta a responsabilidad y complejidad en Sistemas de Suministro de Energía Eléctrica (SSEE), ya que las sobrestimaciones de la demanda real ocasionan sobrecostos en el despacho económico y las subestimaciones provocan imprevisiones de la Reserva Rotante (Regulación Secundaria de Frecuencia) que pueden ocasionar problemas de estabilidad en el caso de contingencias.

El problema del Pronóstico de Demanda que es necesario para la Planificación de la Operación de mediano, corto y muy corto plazo se describe en detalle en esta Investigación. Los modelos y algoritmos desarrollados consideran las incertidumbres asociadas principalmente a variables climáticas como la temperatura, debido a la correlación que tiene en el corto plazo con la demanda de un Sistema Eléctrico.

En esta investigación se resuelve el Pronóstico de Demanda con un horizonte de hasta 168 horas (para la Programación Semanal) incluyendo la actualización diaria mediante la incorporación de nueva información relacionada con las variables explicativas del pronóstico “demanda” (para la Programación Diaria).

Factores climáticos (cambios súbitos en el clima, debido por ejemplo a frentes fríos o lluvias) o aleatorios (paros, huelgas, etc.) que no son considerados en el Pronóstico de Demanda con 1 día de adelanto pueden ser corregidos con el Pronóstico de Demanda de Muy Corto Plazo. Este presenta la ventaja de utilizar la información real del sistema de los últimos periodos horarios de demanda para poder estimar el comportamiento de la nueva tendencia de la demanda del Sistema Eléctrico, disminuyendo así las desviaciones de la demanda real con respecto a la demanda pronosticada (subestimación o sobreestimación).

El concepto presentado en esta Investigación se basa en un modelo no lineal que simula cualitativa y cuantitativamente un sistema multivariable, considerando variables determinísticas (tipo de día) y variables estocásticas (demanda y temperatura), y que acepta sin mayores variaciones distintos horizontes de pronóstico: mediano, corto y muy corto plazo (desde 1 semana hasta 1 hora). Esta modelación Cualitativa – Cuantitativa posee la ventaja de poder incluir n variables explicativas, sin necesidad de una modelación compleja, debido a la formulación multivariable inherente al Razonamiento Inductivo Fuzzy (FIR).

Las entradas del modelo FIR se resuelven a través de la optimización de la función objetivo (error de pronóstico).

Los principales puntos que se resuelven en esta Investigación son:

1. La consideración de un modelo de pronóstico no lineal que, comparado con Modelos Econométricos (ARIMA) y Redes Neuronales Artificiales, logra estimaciones más acertadas de la demanda real del Sistema y con menor incertidumbre asociada (se disminuye la media y la desviación estándar de los errores del pronóstico).
2. El planteamiento de un Modelo Multivariable sin necesidad de complejas modelaciones -como se da en los Modelos ARIMA-, posibilita resolver las Incertidumbres de corto plazo debido a variables climáticas.
3. La obtención de un Modelo Único de Pronóstico válido para distintos horizontes de pronóstico (mediano, corto y muy corto plazo).
4. La determinación óptima de las entradas al Modelo de Pronóstico.

Estos puntos -de crucial importancia en el Pronóstico de cualquier serie temporal-, son tratados con la profundidad necesaria para obtener mejoras esenciales en la disminución del error de pronóstico de demanda con respecto a las Metodologías utilizadas comúnmente en SSEE. La modelación del método propuesto permite además incluir cualquier variable exógena definida por el Planificador sin mayores esfuerzos.

La ventaja de la Metodología FIR radica en que la intervención del Planificador es mínima, ya que los parámetros necesarios del modelo son intuitivos y optimizados por la misma metodología.

Tanto los modelos como los algoritmos desarrollados han sido validados exhaustivamente con datos provenientes de la demanda real del SSEE del Ecuador. Como referencia de comparación se utilizan los pronósticos de demanda obtenidos con Modelos Econométricos – ARIMA por la Dirección de Planeamiento del Operador del Sistema de Ecuador (CENACE).

En esta Investigación se demuestra además que, para encontrar las entradas del modelo de pronóstico, no basta utilizar las metodologías tradicionales para selección de entradas y que están basadas en series temporales (función de autocorrelación y función de autocorrelación parcial) o la Entropía de Shannon utilizada por la Metodología FIR propuesta por Cellier [7]. En efecto, la selección tradicional de entradas en los modelos econométricos se basa en la correlación lineal clásica y -debido a que el pronóstico de demanda es un problema no lineal- no es suficiente esta selección autoregresiva de entradas para seleccionar las variables que realmente influyen sobre el consumo eléctrico. Por lo tanto, es necesario utilizar alguna herramienta de correlación no lineal, ya que a veces variables con una alta correlación no lineal, respecto a la variable estudiada, se descartan por que la correlación lineal no es capaz de medir esa relación.

Frente a la limitación que presentan las metodologías tradicionalmente usadas en series de tiempo, las cuales encuentran mediante la función de autocorrelación parcial un grupo de entradas que son validas para un determinado periodo de tiempo (por ejemplo una estación del año), la nueva metodología aquí propuesta realiza un proceso dinámico de selección de entradas del modelo, que posibilita realizar un seguimiento a la tendencia de la demanda en forma instantánea.

Además, estos desarrollos poseen un potencial de aplicación que excede en gran medida el ámbito de la presente Investigación y son por tanto aplicables a otros problemas de la Planificación y Operación de sistemas de potencia como por ejemplo el Pronóstico de Recurso Hídrico. En general, este Modelo es válido para la modelación de cualquier Sistema Multivariable que esté representando un fenómeno físico.

Por estas razones, tanto los modelos y algoritmos desarrollados en esta Investigación como los resultados que con ellos se obtienen en la aplicación práctica a sistemas reales son sin duda un aporte significativo en el área de la Planificación de la Operación de SSEE.

APÉNDICES

A. ALGORITMO PARA DETERMINAR EL NÚMERO DE CLASES

Este algoritmo es propuesto por Chiu [13].

Algoritmo

Si se considera n puntos $\{x_1, \dots, x_n\}$ pertenecientes a un espacio L -dimensional y cada punto es candidato a centro del cluster (clase). Se define la medida de densidad de un punto x_i con la ecuación (A.1).

$$D_i^0 = \sum_{j=1}^n e^{\left(-\frac{\|x_i - x_j\|^2}{(r_a/2)^2} \right)} \quad (\text{A.1})$$

donde; r_a : es una constante positiva

Un punto tendrá un valor grande de densidad D^0 si tiene algunos vecinos, donde el radio r_a define su vecindad.

Una vez que la densidad es calculada, se selecciona como centro de la clase (x_{c1}) el dato con mayor D y se vuelve a calcular la densidad para cada punto como se muestra en la ecuación (A.2).

$$D_i^1 = D_i^0 - D_{c1} e^{\left(-\frac{\|x_i - x_{c1}\|^2}{(r_b/2)^2} \right)} \quad (\text{A.1})$$

donde; r_b : es una constante positiva

Se observa que en los datos cercanos al primer centro de la clase (x_{c1}) se disminuye significativamente su densidad. La constante r_b define otra vecindad mide la reducción de la densidad inicial y normalmente es $r_b = 1.5 * r_a$.

Se repite el cálculo de la densidad tomando como centro de la clase (x_{c2}), el dato con mayor D^1 . Este proceso se repite hasta que se genera un número suficiente de clases.

Este método es proporcional al número de datos e independiente de la dimensión del problema a considerar.

B. ALGORITMOS PARA DETERMINAR LOS LÍMITES DE LAS FUNCIONES DE MEMBRESÍA

B.1. Equal Frequency Partition

Se basa en la idea de que la expresividad (medida de la información que el modelo provee) será maximizada si cada clase se observada igual número de veces.

Algoritmo

- 1) La evolución del valor observado de cada variable es clasificada dentro de un orden ascendente.
- 2) El vector ordenado se divide en k (número de clases) intervalos de igual número de elementos.
- 3) Los límites para una determinada clase, se eligen en cualquier parte de los valores extremos del respectivo intervalo, por ejemplo usando la media aritmética.

B.2. K-Means

A este algoritmo se lo conoce también como C-means. Este algoritmo clasifica un vector x_j ($j = 1, \dots, n$) en c clases G_i ($i = 1, \dots, c$) y encuentra el centro para cada clase tal que una función de costo (función objetivo) es minimizada. Si se utiliza la distancia euclidiana como medición entre el vector x_k del grupo i y el correspondiente centro c_i , la función de costo se define mediante la ecuación (B.1).

$$J = \sum_{i=1}^c J_i = \sum_{i=1}^c \left(\sum_{k, x_k \in G_i} \|x_k - c_i\|^2 \right) \quad (\text{B.1})$$

donde; J_i : es la función de costo i dentro de una clase i

El valor de J_i depende de las propiedades geométricas de G_i y de la ubicación de c_i .

Las particiones son definidas por una matriz de membresía U [$c \times n$], donde el elemento u_{ij} es 1 si el valor x_j pertenece a la clase i y 0 si no. Una vez que los centros de la clase c_i están fijos, de la ecuación (B.1) se pueden derivar los valores u_{ij} mediante la ecuación (B.2).

$$u_{ij} = \begin{cases} 1 & \text{si } \|x_j - c_i\|^2 \leq \|x_j - c_k\|^2 \ (k \neq i) \\ 0 & \text{resto} \end{cases} \quad (\text{B.2})$$

Entonces un punto cual quiera puede estar solamente en una clase y la matriz de membresía U tiene las siguientes propiedades (ecuación B.3).

$$\sum_{i=1}^c u_{ij} = 1, \forall j = 1, \dots, n \Rightarrow \sum_{i=1}^c \sum_{j=1}^n u_{ij} = n \quad (\text{B.3})$$

Por otro lado, si la membresía u_{ij} es fija, entonces el óptimo centro c_i que minimiza la ecuación (B.1) es la media (ecuación (B.4)) de todos los vectores pertenecientes al grupo i .

$$c_i = \frac{1}{|G_i|} \sum_{k, x_k \in G_i} x_k \quad (\text{B.4})$$

donde; $|G_i| = \sum u_{ij} \ (j=1, \dots, n)$

Algoritmo

- 1) Se inicializa aleatoriamente de entre el conjunto de datos a los i centros de cada cluster c_i .
- 2) Se determina la matriz de membresía U utilizando la ecuación (B.2).
- 3) Se calcula la función de costo mediante la ecuación (B.1) y como condiciones de parada del algoritmo se tiene: un criterio de tolerancia o un valor umbral respecto de la iteración anterior.
- 4) Se actualiza los centros de cada cluster con la ecuación (B.4). Se regresa al punto 2).

El algoritmo es iterativo y no se garantiza que converja a una óptima solución. La performance de este algoritmo depende de la posición inicial de cada centro c_i , por lo que se recomienda correr varias veces el algoritmo y en cada corrida se utiliza diferentes posiciones iniciales.

B.4. Fuzzy C-Means

También conocido como “fuzzy Isodata”. En este algoritmo cada punto pertenece a un cluster para un grado de membresía. Besdek propuso este algoritmo [3], mejorando los anteriores algoritmos de clustering descritos previamente.

Este algoritmo particiona un vector $x_i \ (i = 1, \dots, n)$ en c grupos fuzzy y encuentra el centro para cada clase tal que la función de costo es minimizada. La diferencia con respecto al anterior algoritmo, se basa en el concepto de que un dato puede pertenecer a varios grupos con un grado de pertenencia especificado por una función de membresía continua entre $[0,1]$, por lo que la matriz U puede tener elementos cuyos valores pueden variar continuamente entre 0 y 1. Además para un grupo de datos, se restringe la sumatoria de los grados de pertenencia a la unidad, mediante la ecuación (B.5).

$$\sum_{i=1}^c u_{ij} = 1, \forall j = 1, \dots, n \quad (\text{B.5})$$

La función objetivo para este algoritmo (ecuación (B.6)) es una generalización de la ecuación (B.1).

$$J(U, c_1, \dots, c_n) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 \quad (\text{B.6})$$

donde; c_j : centro del grupo fuzzy i
 d_i : es la distancia euclidiana
 m : exponente de ponderación $[1, \infty)$

Para encontrar las condiciones necesarias para que la ecuación (B.6) alcance un mínimo, procedemos a transformar la función objetivo (ecuación (B.7)).

$$\begin{aligned} \bar{J}(U, c_1, \dots, c_n, \lambda_1, \dots, \lambda_n) &= \bar{J}(U, c_1, \dots, c_n) + \sum_{j=1}^n \lambda_j \left(\sum_{i=1}^c u_{ij} - 1 \right) \\ &= \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 + \sum_{j=1}^n \lambda_j \sum_{i=1}^c \lambda_j \left(\sum_{i=1}^c u_{ij} - 1 \right) \end{aligned} \quad (\text{B.7})$$

donde; λ_j : ($j = 1, \dots, n$) multiplicadores de Lagrange de la ecuación (B.5)

Para alcanzar el mínimo de la ecuación (B.5), se procede a diferenciar con respecto a sus argumentos y se encuentra los valores de c y u (ecuación (B.8)).

$$c_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad \text{y} \quad u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{jk}} \right)^{\frac{2}{m-1}}} \quad (\text{B.8})$$

La ecuación (B.8) son las condiciones necesarias para alcanzar el mínimo.

Algoritmo

- 1) Se inicializa la matriz de membresía U con valores aleatorios entre 0 y 1, tal que se debe satisfacer la restricción de la ecuación (B.4).
- 2) Se calculan c centros de cada clase c_i usando la ecuación (B.7).
- 3) Calcular la función de costo con la ecuación (B.5). El algoritmo se detiene si: la función de costo es menor que una tolerancia dada o si la diferencia con respecto a la iteración anterior es menor que un umbral definido.
- 4) Se calcula una nueva matriz U usando la ecuación (B.7). Se regresa al paso 2).

En este algoritmo tampoco se asegura una convergencia a la óptima solución, ya que depende de la inicialización de los centros de los clusters, por lo que también se recomienda correr varias veces el algoritmo y en cada corrida utilizar diferentes centros en cada cluster.

C: ALGORITMO DE REBOTES SIMULADOS (SRA)

C.1. Algoritmos Evolutivos

Los algoritmos evolutivos (AE) son algoritmos de aproximación basados en analogías con procesos evolutivos de la naturaleza. Estrictamente, se entiende por AE a aquellos que utilizan los principios de la teoría de la evolución: selección, mutación y recombinación. Sin embargo, estas características se presentan también en procesos físicos tales como el proceso de cristalización de los metales en el recocido, el proceso de adsorción y las colisiones inelásticas por citar algunos ejemplos.

Los algoritmos evolutivos constan de los siguientes elementos:

- un conjunto de soluciones iniciales,
- operadores de transición hacia nuevas soluciones,
- operadores de control de aceptación de nuevas soluciones,
- una función objetivo (a optimizar).

Dependiendo de estos cuatro elementos, encontramos algoritmos que operan simultáneamente sobre varias soluciones (Genetic Algorithms) o sobre una solución a la vez (Simulated Annealing, Simulated Rebounding), utilizando diferentes operaciones de transición (probabilísticas o determinísticas) hacia una o varias soluciones según sea el caso, que son aceptadas o rechazadas según un criterio (probabilístico o determinístico) que depende del parámetro de control y la función objetivo.

Las técnicas de optimización basadas en estos principios [33] son:

- Algoritmos Genéticos (AG)
- Estrategias Evolutivas (EE)
- Programación Genética

Los AE forman parte de un campo mayor en el universo científico, cual es la inteligencia artificial (IA) y son utilizados a menudo en conjunción con otras técnicas de IA: Sistemas Borrosos, Redes Neuronales Artificiales, Sistemas Expertos y Razonamiento Inductivo Fuzzy. Este campo de la ciencia también es llamado Inteligencia computacional (Soft Computing).

Las estrategias evolutivas (EE) fueron concebidas en 1960 por Rechenberg para resolver problemas de optimización paramétrica. Las EE emplean una codificación real de las soluciones, es decir, mediante vectores con números reales donde cada elemento representa un parámetro a optimizar; utilizan el operador de mutación como operador de búsqueda o transición hacia nuevas soluciones y una población unitaria (es decir, operan

sobre una solución por vez). Dentro de este campo pueden incluirse los algoritmos de Annealing y Rebounding, como así también en forma restringida la Programación Dinámica. Las EE pueden considerarse tanto como una abstracción de la evolución genética a nivel individual, o como una analogía con procesos evolutivos de otro tipo que el genético, principalmente físico-químicos. De estas analogías surgen diferentes operadores de control de aceptación de nuevas soluciones.

C.2. Búsqueda Local Aleatoria

Para evitar la suboptimalidad inherente a los algoritmos de búsqueda local (BL), se suele acudir a los siguientes recursos:

- ejecución de BL para un número grande de soluciones iniciales,
- introducción de estructuras de vecindad complejas, de manera de asegurar una búsqueda eficiente en el espacio de soluciones y
- aceptar, en forma controlada, transiciones hacia soluciones “peores”, con el objetivo de evitar caer en un mínimo local.

El último recurso puede implementarse mediante un algoritmo o heurística controlada que acepta, según una probabilidad dada, transiciones o soluciones que se alejan del óptimo. Esto hace que la búsqueda local tenga un comportamiento aleatorio y que para una misma instancia del problema y una misma solución inicial, las soluciones encontradas sean a veces diferentes. Los algoritmos más eficientes que se basan en esta filosofía son: Simulated Annealing y Simulated Rebounding. Estos algoritmos se basan en analogías de procesos físicos.

En la investigación realizada por A. Hoese [24] se concluye que el algoritmo SRA llega a muy buenos resultados en menor tiempo de cálculo que el utilizado por el algoritmo SAA. El SRA, sin embargo, obtiene soluciones que dependen en cierto grado de las condiciones iniciales del problema.

Comparaciones realizadas con AG [44] muestran que el SRA requiere aproximadamente del 25% de tiempo de cálculo necesario para obtener soluciones de la misma calidad que con AG en problemas de dimensión $d=4$ (similar a nuestra investigación). Esto permitiría realizar hasta 4 corridas del SRA a fin de minimizar la dependencia de la solución respecto de la solución de partida, con lo cual los tiempos de cálculo se podrían comparar con los de AG.

El SRA fue presentado en 1993 por A. Hoese [22] como una heurística evolutiva basada en los mismos principios que el Algoritmo de Recocido Simulado (Simulated Annealing Algorithm – SAA) pero haciendo analogía con el proceso de la física cinética del choque inelástico de un cuerpo sometido a la acción de un campo gravitatorio. El algoritmo fue diseñado en principio para un tipo específico de problemas de optimización combinatorial paramétrica, donde el número de parámetros constituyentes de la solución no es muy elevado y la función de costo requiere de un esfuerzo computacional importante para obtener el valor de costo de cada solución (obtenida en estos casos mediante métodos simulativos). Es estos problemas, la velocidad de convergencia es fundamental, y la búsqueda sobre el espacio de soluciones debe ser

eficiente, en el sentido de que las transiciones propuestas sean las mínimas indispensables para alcanzar una solución cercana al óptimo global.

C.3. La simulación del proceso físico de rebotes inelásticos

Suponiendo que se lanza una pelota con coeficiente de elasticidad cercano a 1 sobre una superficie S , la pelota caerá sobre la superficie debido a la acción del campo gravitatorio en el cual esta inmersa. El choque producido por la caída podrá afectar a la energía cinética de la pelota si parte de esta energía es disipada en deformaciones permanentes (choque inelástico). Estas deformaciones serán proporcionales a la energía cinética del choque, siendo esta la equivalente a la diferencia de energía potencial desde el punto de choque hasta la altura desde la cual cae la pelota. Si el proceso no es detenido, los rebotes se detendrán cuando la energía cinética de la pelota sea cero. En este proceso de rebotes inelásticos, la pelota choca con distintos puntos de la superficie, desviada por los mismos rebotes sobre una superficie no plana. Si la superficie posee “valles” y “lomas”, los puntos pertenecientes a un valle tendrán menor energía potencial que las lomas. Por la acción de la gravedad, la pelota buscará detenerse sobre algún valle, y dependiendo de la energía cinética que posea, será capaz de saltar montes para explorar nuevos valles.

La analogía con un problema de optimización combinatorial se realiza haciendo una equivalencia entre la superficie de rebotes y el espacio de soluciones, y entre la altura de cada punto de la superficie y el valor de la función de costo correspondiente a esa solución. El SRA resulta así una extrapolación multidimensional del proceso físico de rebotes. Debido a la acción de la gravedad y a la pérdida gradual de energía cinética y potencial debido a las colisiones, al final del proceso de rebotes simulados, la pelota se detiene sobre el mínimo global (punto de menor altura de la superficie) o sobre algún mínimo local cercano al global. La Figura C.1 muestra por simplicidad un ejemplo unidimensional: como se muestra con las líneas de trayectoria, la pelota puede alcanzar el mínimo global (escalón 11) si es lanzada desde el punto A o C; sin embargo es atrapada en el mínimo local del escalón 14 si la pelota es lanzada desde el punto B. Es decir, el mínimo global es alcanzado si:

- la energía potencial inicial (E_0), y
- el porcentaje de pérdida de energía debido a colisiones inelásticas (α)

permiten saltar todos los mínimos locales durante el proceso de rebotes.

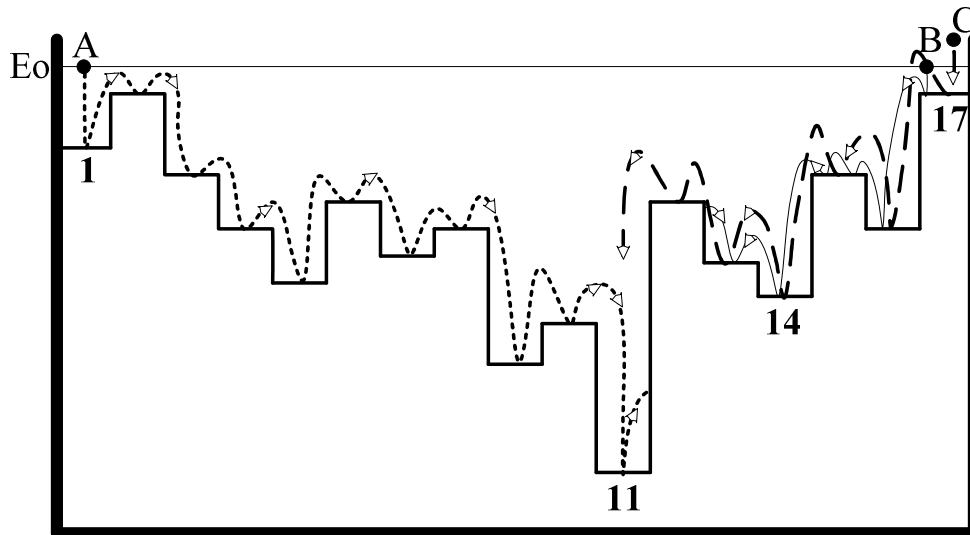


Figura C.1. Ejemplo unidimensional de rebotes simulados

En el SRA las deformaciones producidas por las colisiones pueden ser de tres tipos:

Deformación transitoria, es decir, que la pelota no pierde energía cinética por ser la altura de la caída pequeña.

Deformación leve, cuando la pelota rebota sobre las paredes de un valle por ser su energía potencial menor que la altura del punto de destino, disminuyendo su energía potencial en un valor pequeño.

Deformación permanente, cuando la pelota pierde energía cinética debido a una colisión provocada por una altura considerable de caída.

Estos tipos de deformación actúan en el proceso de optimización de diferente manera. La deformación transitoria es la que permite explorar el espacio de búsqueda. Si sólo existiera este tipo de deformaciones, el proceso no se detendría nunca, si no fuera por la pérdida de energía provocada por fricción y rozamiento. La deformación leve disminuye la energía de la pelota cuando es atrapada en un valle, y provoca la detención del proceso si no existe ninguna solución con energía potencial menor que la de la pelota (valle profundo). La deformación permanente acelera la convergencia hacia el óptimo, ya que disminuye la energía potencial de la pelota cuando ésta es muy superior a la de la superficie en el entorno de rebotes.

Para la i -ésima colisión inelástica (deformación permanente), el valor de la energía de la pelota es obtenida mediante la ecuación (C.1).

$$E_i = \alpha E_{i-1}; \quad 0 < \alpha < 1 \quad (\text{C.1})$$

donde; α : tiene su analogía en el proceso físico con el **coeficiente de elasticidad** de la pelota. Para un material totalmente elástico ($\alpha=1$), no existen deformaciones permanentes y por lo tanto la energía de la pelota no disminuye con las colisiones.

C.4. Estructura de vecindad del SRA

Las transiciones son hechas en una estructura de vecindad Φ_s alrededor del punto de rebote s , la cual puede variar de un problema a otro. Por simplicidad en el ejemplo unidimensional de la figura (C.1), la pelota sólo puede rebotar a escalones vecinos (a izquierda o derecha). Es lo que se define como **mínima estructura de vecindad Φ** , cuyo módulo puede calcularse para el caso d-dimensional mediante la ecuación (C.2).

$$|\Phi| = 2 * \dim(\Theta) = 2 * d \quad (C.2)$$

Una estructura de vecindad Φ_s define las soluciones que se encuentran alrededor de una solución determinada s , es decir, las soluciones que se encuentran adyacentes a ella. Una posible definición se muestra en la ecuación (C.3).

$$z \in \Phi_s \Leftrightarrow \begin{cases} x_{j(z)} = x_{j(s)} \pm n & (j=1, \dots, m) \\ x_{i(z)} = x_{i(s)} & \forall (i \neq j) \end{cases} \quad (s \in \mathfrak{S}^m) \quad (C.3)$$

donde; n : número aleatorio con distribución uniforme

Por lo que z es una solución vecina de s y $x_{j(s)}$ es el j -ésimo elemento de la solución. La estructura de vecindad y el mecanismo de generación para las soluciones vecinas debe ser adaptado al espacio de soluciones del problema dado, es decir el mecanismo de generación de la ecuación (C.3) no es apropiado para espacio de soluciones pertenecientes a los reales.

El procedimiento a seguir para la construcción de una solución vecina sería:

- elegir un j entre 1 y m en forma aleatoria con probabilidad uniforme (elección de la variable a variar),
- elegir aleatoriamente un entero n y
- variar el $x_{j(s)}$ según el valor n obtenido.

La dimensión de tal estructura de vecindad será: $\Theta=m*n$, donde m es la dimensión de la configuración v .

C.5. Criterio de aceptación y detención del SRA

El modulo de Φ define el número de soluciones pertenecientes a la mínima estructura de vecindad. La solución de transición es elegida aleatoriamente entre las pertenecientes a la estructura de vecindad y si la energía correspondiente a la nueva solución es menor que la energía del proceso, es aceptada como nuevo punto de rebote. Caso contrario, es elegida una nueva solución correspondiente a la estructura de vecindad. El criterio de aceptación del SRA se muestra en la ecuación (C.4).

$$\Pr(j = \text{próxima_solución} /_k) \begin{cases} 1 \Rightarrow f_{(j)} \leq E_k \\ 0 \Rightarrow f_{(j)} > E_k \end{cases} \quad (C.4)$$

El proceso finaliza cuando la energía del proceso no es suficiente para alcanzar una nueva solución, luego de pruebas sobre la totalidad de la estructura de vecindad de la solución i , con lo cual el criterio de detención será: $f_{(x)} > E_k$ para todo x elemento de V_i .

En la práctica, el criterio de detención se obtiene reduciendo la energía de la pelota cada vez que un rebote no permite alcanzar una solución vecina (es decir, una solución vecina cuyo valor de costo es mayor que la energía de la pelota). Esta reducción por choques frustrados, que en analogía con el proceso físico se corresponde con las deformaciones leves, debe sin embargo permitir la exploración completa del espacio de soluciones. De aquí que la reducción de la energía cinética por choques frustrados β viene dada por la ecuación (C.5).

$$\beta < \frac{E_i^o - f_{(i)}}{|\Phi|} \quad (\text{C.5})$$

Y por lo tanto la reducción de la energía del proceso con cada choque frustrado k se expresa mediante la ecuación (C.6).

$$E_i^{k+1} = E_i^k - \beta; \quad k = 0, 1, \dots, |\Phi| - 1 \quad (\text{C.6})$$

donde; E_i^{k+1} : es la energía de la pelota en la i -ésima iteración luego del k -ésimo rebote frustrado.

$f_{(i)}$: energía potencial de la solución alcanzada en la i -ésima iteración, es decir, el valor de la función de costo para la solución del punto de rebote.

Esta reducción por choques frustrados detiene el proceso de rebotes en la solución alcanzada en la i -ésima iteración una vez que intenta explorar completamente el espacio de soluciones vecino sin encontrar una solución factible de ser alcanzada.

C.6. Implementación práctica del SRA

Para la implementación práctica del SRA es necesario contar con:

- una condición inicial (E_o y la variante o solución inicial de búsqueda ν),
- el valor del parámetro (porcentaje de conservación de energía),
- una estructura de vecindad V (estructura de variantes cercanas o adyacentes a la evaluada),
- un mecanismo de generación para la determinación de la próxima variante a ser evaluada.

El algoritmo no es robusto, ya que intuitivamente se ve que la calidad de la solución encontrada (óptimo local cercano al global) depende en general de la solución inicial adoptada. Como contraparte, es un algoritmo muy eficiente en cuanto a tiempo de cálculo y sólo necesita evaluar una porción muy reducida del espacio de soluciones. Debido a estas características, es aconsejable realizar la optimización con distintas soluciones iniciales, con el fin de:

- asegurar una solución óptima global o una muy cercana a esta,
- encontrar un conjunto de soluciones subóptimas muy cercanas a la global.

Puesto que el conjunto de soluciones subóptimas se encuentran analizando la trayectoria seguida por el algoritmo para encontrar la solución óptima (espacio de soluciones evaluado), y con el fin de que este conjunto sea lo más representativo del espacio de soluciones total Θ , normalmente se consideran un número de soluciones iniciales ν igual a los vértices del espacio de solución Θ . Se debe notar que una condición inicial podría no pertenecer al CFMC (Θ_N), en cuyo caso se aplica la estructura de vecindad a la condición inicial respectiva.

El valor inicial de energía E_0 de la solución inicial puede adoptarse en la práctica [23] como una función del parámetro α y el valor de la función de costo de la solución inicial $f(v)$ viene dado por la ecuación (C.7).

$$E_0 = \frac{f(v)}{\alpha} \quad (\text{C.6})$$

donde; $0,7 < \alpha < 0,99$

El SRA se aplica varias veces, cada vez con una solución inicial distinta.

El pseudocódigo del SRA se muestra a continuación:

solución_rebounding (solución x , real α , real d)

```

{
  solución vn
  real  $\beta$ ,  $\delta$ , Eo, E1, F_v, F_vn

  F_v=PDPCP(v);
  Eo=F_v /  $\alpha$ ; // energía inicial de la pelota
   $\beta$ =(Eo-F_v) / 1,1 / d; // decremento por choques frustrados
   $\delta$ =(9+  $\alpha$ ) / 10; // decremento por rozamiento

  do
  {
    E1= $\alpha$  * Eo; // próximo nivel de energía (colisión inelástica)
    do
    {
      vn=vecino_de(v);
      F_vn=PDPCP(vn);
      if ( $\delta$ *Eo < F_vn) // deformación leve (rebote frustrado)
      {
        Eo=Eo- $\beta$ ;
      }
      else if (E1 ≤ F_vn) // deformación transitoria (choque elástico)
      {
        v=vn;
        F_v=F_vn;
         $\beta$ =(Eo-F_v) / 1,1 / d;
      }
      else // deformación permanente (choque inelástico)
      {
        Eo=E1;
        v=vn;
        F_v=F_vn;
         $\beta$ =(Eo-F_v) / 1,1 / d;
        break do; // recalculamos para el nuevo nivel de energía
      }
    } while (F_v <= Eo);
  } while (F_v <= Eo);

  return (v); // retorna la solución encontrada
}

```

D: RESULTADOS DE LAS SIMULACIONES

En este anexo se realiza un análisis de las distintas dimensiones de ν . En las tablas se detalla tanto el error del grupo de datos de entrenamiento ($E_{(t)}$) como el error promedio del grupo de validación o test ($E_{TEST(t)}$) y el tiempo de simulación en minutos.

D.1. PDCP para la dimensión $\nu_3 - 1$ entradas para la variable demanda

Los resultados para una profundidad de 25 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t-16),F(t-8)	8,9415	11,1297	3,8
	A(t-5),E(t-17),F(t-24)	3,5691	3,7410	
2	A(t-7),E(t-11),F(t-15)	9,2062	10,1939	2,5
	A(t),E(t-1),F(t-24)	3,1639	3,4662	

Tabla D.1. PDCP con 1 entrada para la demanda y con una profundidad de 25

Los resultados para una profundidad de 49 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-4),E(t-17),F(t-24)	3,5551	3,9484	4,7
	A(t-4),E(t-46),F(t-48)	6,6530	5,4523	
2	A(t),E(t-1),F(t-24)	3,3307	3,4789	5,3
	A(t-5),E(t-23),F(t-48)	5,1947	5,8430	

Tabla D.2. PDCP con 1 entrada para la demanda y con una profundidad de 49

Los resultados para una profundidad de 73 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-4),E(t),F(t-24)	3,6597	3,8713	6,9
	A(t-2),E(t-64),F(t-72)	7,3836	6,9794	
2	A(t),E(t-62),F(t-24)	3,1623	3,7025	7,4
	A(t-5),E(t-19),F(t-72)	6,5345	7,4717	

Tabla D.3. PDCP con 1 entrada para la demanda y con una profundidad de 73

La máscara con la que se obtiene el menor error de validación tiene de profundidad 25. Un resumen para las 2 semanas de test se muestra en la Tabla D.4.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1)$	$E_{(t)}$	$E_{TEST(t)}$
1	A(t-5),E(t-17),F(t-24)	3,5691	3,7410
2	A(t),E(t-1),F(t-24)	3,1639	3,4662

Tabla D.4. PDCP con 1 entrada para la demanda

D.2. PDCP para la dimensión v_4 - 2 entradas para la variable demanda

Los resultados para una profundidad de 25 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t),E(t-11),F(t-9),F(t-5)	3,3563	4,1001	7,8
	A(t),E(t-1),F(t-24),F(t-1)	2,4247	3,0145	
	A(t-5),F(t),F(t-24),F(t-10)	1,8284	1,8621	
2	A(t-2),E(t-13),F(t-9),F(t-5)	3,7895	5,1808	6,8
	A(t-6),E(t-2),F(t-24),F(t-12)	1,6486	1,7821	
	A(t-6),E(t-23),F(t-24),F(t-11)	1,5859	1,9033	

Tabla D.5. PDCP con 2 entradas para la demanda y con una profundidad de 25

Los resultados para una profundidad de 49 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t),E(t-8),F(t-9),F(t-5)	3,3356	4,1211	8,3
	A(t-5),E(t-18),F(t-48),F(t-25)	2,1556	1,9027	
	A(t-5),E(t-29),F(t-42),F(t-24)	1,7405	1,5928	
2	A(t),E(t-3),F(t-10),F(t-1)	2,6680	3,9811	9,3
	A(t-5),E(t-27),F(t-47),F(t-24)	1,4133	1,6627	
	A(t-5),E(t-9),F(t-48),F(t-40)	1,6314	2,2583	

Tabla D.6. PDCP con 2 entradas para la demanda y con una profundidad de 49

Los resultados para una profundidad de 73 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t),E(t-30),F(t-9),F(t-1)	3,4147	7,2289	10,9
	A(t-5),E(t-62),F(t-72),F(t-10)	2,6166	2,6807	
	A(t-5),E(t-37),F(t-71),F(t-48)	2,5621	2,7093	
2	A(t),E(t-3),F(t-10),F(t-1)	2,6719	4,0409	12,4
	A(t-4),E(t-72),F(t-71),F(t-24)	1,4721	1,8845	
	A(t-5),E(t-24),F(t-71),F(t-48)	1,6707	2,1971	

Tabla D.7. PDCP con 2 entradas para la demanda y con una profundidad de 73

La máscara con la que se obtiene el menor error de validación tiene de profundidad 49. Un resumen para las 2 semanas de test se muestra en la Tabla D.8.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2)$	$E_{(t)}$	$E_{TEST(t)}$
1	A(t-5),E(t-29),F(t-42),F(t-24)	1,7405	1,5928
2	A(t-5),E(t-27),F(t-47),F(t-24)	1,4133	1,6627

Tabla D.8. PDCP con 2 entradas para la demanda

D.3. PDCP para la dimensión v_5 - 3 entradas para la variable demanda

Los resultados para una profundidad de 25 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t),E(t-6),F(t-7),F(t-2),F(t-1)	1,5932	2,0695	20,4
	A(t-5),E(t-3),F(t-24),F(t-10),F(t-3)	1,3434	1,2827	
	A(t-4),E(t-23),F(t-24),F(t-18),F(t-1)	1,3387	1,1753	
	A(t-5),E(t-0),F(t-24),F(t-22),F(t-19)	1,6345	1,9765	
2	A(t-1),E(t-7),F(t-12),F(t-9),F(t-1)	1,3080	1,8652	20,5
	A(t-3),E(t-3),F(t-24),F(t-17),F(t-1)	1,0145	1,4659	
	A(t-4),E(t-12),F(t-24),F(t-17),F(t-1)	1,0437	1,4734	
	A(t-4),E(t-12),F(t-24),F(t-17),F(t-1)	1,0437	1,4734	

Tabla D.9. PDCP con 3 entradas para la demanda y con una profundidad de 25

Los resultados para una profundidad de 49 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-1),E(t-23),F(t-23),F(t-11),F(t-1)	1,4694	1,2504	23,2
	A(t-4),E(t-9),F(t-48),F(t-7),F(t-3)	1,6450	1,3433	
	A(t-4),E(t-24),F(t-48),F(t-41),F(t-1)	1,5237	1,3216	
	A(t-1),E(t-25),F(t-47),F(t-43),F(t-24)	1,5789	1,3326	
2	A(t),E(t-35),F(t-24),F(t-11),F(t-1)	1,1643	1,5369	33,4
	A(t-4),E(t-9),F(t-46),F(t-10),F(t-1)	1,2079	1,5097	
	A(t-5),E(t-4),F(t-48),F(t-39),F(t-1)	1,2189	1,6855	
	A(t-5),E(t-25),F(t-48),F(t-41),F(t-35)	1,1942	1,7185	

Tabla D.10. PDCP con 3 entradas para la demanda y con una profundidad de 49

Los resultados para una profundidad de 73 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-4),E(t-23),F(t-24),F(t-18),F(t-1)	1,3242	1,2054	33,9
	A(t),E(t),F(t-69),F(t-23),F(t-1)	1,4892	1,3442	
	A(t),E(t-37),F(t-72),F(t-45),F(t-24)	1,5641	1,3844	
	A(t-4),E(t-7),F(t-72),F(t-69),F(t-48)	1,7560	1,4495	
2	A(t),E(t-35),F(t-24),F(t-11),F(t-1)	11682	15343	35,1
	A(t),E(t-22),F(t-72),F(t-24),F(t-21)	1,1371	1,6782	
	A(t),E(t-26),F(t-57),F(t-47),F(t-24)	1,1569	1,5011	
	A(t-4),E(t-37),F(t-71),F(t-58),F(t-48)	1,2750	1,7990	

Tabla D.11. PDCP con 3 entradas para la demanda y con una profundidad de 73

La máscara con la que se obtiene el menor error de validación tiene de profundidad 25. Un resumen para las 2 semanas de test se muestra en la Tabla D.12.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3)$	$E_{(t)}$	$E_{TEST(t)}$
1	A(t-4),E(t-23),F(t-24),F(t-18),F(t-1)	1,3387	1,1753
2	A(t-3),E(t-3),F(t-24),F(t-17),F(t-1)	1,0145	1,4659

Tabla D.12. PDCP con 3 entradas para la demanda

D.4. PDCP para la dimensión v_6 - 4 entradas para la variable demanda

Los resultados para una profundidad de 25 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$	t [min]
1	A(t-4),E(t-9),F(t-22),F(t-3),F(t-2),F(t-1)	1,4889	1,3786	33,6
	A(t-5),E(t-18),F(t-23),F(t-13),F(t-7),F(t-1)	1,2644	1,2141	
	A(t-5),E(t-1),F(t-24),F(t-17),F(t-4),F(t-1)	1,2818	1,0390	
	A(t-4),E(t-5),F(t-24),F(t-23),F(t-12),F(t-2)	1,3094	1,1855	
	A(t-5),E(t-10),F(t-24),F(t-22),F(t-18),F(t-1)	1,2898	1,1209	
2	A(t-1),E(t-15),F(t-16),F(t-8),F(t-3),F(t-1)	1,1405	1,6443	45,4
	A(t-4),E(t),F(t-24),F(t-17),F(t-7),F(t-1)	1,0134	1,4512	
	A(t),E(t-10),F(t-23),F(t-21),F(t-17),F(t-1)	1,0004	1,3537	
	A(t-4),E(t-22),F(t-24),F(t-18),F(t-17),F(t-1)	0,9655	1,3909	
	A(t),E(t),F(t-24),F(t-23),F(t-19),F(t-14)	1,0962	1,4316	

Tabla D.13. PDCP con 4 entradas para la demanda y con una profundidad de 25

Los resultados para una profundidad de 49 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$	t [min]
1	A(t-5),E(t-6),F(t-9),F(t-6),F(t-2),F(t-1)	1,5391	1,3876	42,7
	A(t-5),E(t-13),F(t-22),F(t-16),F(t-6),F(t-1)	1,2958	1,1905	
	A(t-5),E(t-6),F(t-44),F(t-23),F(t-3),F(t-1)	1,2883	1,3851	
	A(t-4),E(t),F(t-45),F(t-37),F(t-23),F(t-1)	1,4304	1,2248	
	A(t-3),E(t-15),F(t-48),F(t-41),F(t-22),F(t-1)	1,3957	1,2237	
2	A(t-1),E(t),F(t-24),F(t-17),F(t-12),F(t-1)	0,9951	1,3854	45,7
	A(t),E(t-31),F(t-44),F(t-16),F(t-7),F(t-1)	1,0996	1,5793	
	A(t-4),E(t-11),F(t-48),F(t-40),F(t-8),F(t-1)	1,1074	1,5140	
	A(t-1),E(t-39),F(t-48),F(t-41),F(t-35),F(t-1)	1,1061	1,5512	
	A(t-5),E(t-12),F(t-48),F(t-42),F(t-41),F(t-35)	1,1728	1,6027	

Tabla D.14. PDCP con 4 entradas para la demanda y con una profundidad de 49

Los resultados para una profundidad de 73 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$	t [min]
1	A(t-5),E(t-15),F(t-22),F(t-16),F(t-6),F(t-1)	1,2898	1,2670	47,7
	A(t-8),E(t-4),F(t-52),F(t-23),F(t-22),F(t-1)	1,4190	1,2316	
	A(t-1),E(t),F(t-72),F(t-70),F(t-23),F(t-1)	1,5211	1,2126	
	A(t),E(t-56),F(t-72),F(t-69),F(t-25),F(t-23)	1,4814	1,2668	
	A(t-5),E(t-4),F(t-72),F(t-66),F(t-48),F(t-46)	1,5707	1,4779	
2	A(t),E(t-3),F(t-24),F(t-21),F(t-19),F(t-1)	1,0213	1,3567	51,8
	A(t-3),E(t-18),F(t-71),F(t-24),F(t-18),F(t-2)	1,0142	1,4328	
	A(t-4),E(t-12),F(t-56),F(t-36),F(t-23),F(t-1)	1,0490	1,4758	
	A(t-5),E(t-21),F(t-72),F(t-71),F(t-36),F(t-24)	1,1066	1,5151	
	A(t-4),E(t-72),F(t-72),F(t-71),F(t-58),F(t-48)	1,2417	1,7335	

Tabla D.15. PDCP con 4 entradas para la demanda y con una profundidad de 73

La máscara con la que se obtiene el menor error de validación tiene de profundidad 25. Un resumen para las 2 semanas de test se muestra en la Tabla D.16.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$
1	A(t-5),E(t-18),F(t-23),F(t-13),F(t-7),F(t-1)	1,2644	1,2141
2	A(t-4),E(t-22),F(t-24),F(t-18),F(t-17),F(t-1)	0,9655	1,3909

Tabla D.16. PDCP con 4 entradas para la demanda

D.5. PDCP para la dimensión v_7 - 5 entradas para la variable demanda

Los resultados para una profundidad de 25 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5)$	$E_{(t)}$	$E_{TEST(t)}$	t[<i>min</i>]
1	A(t-5),E(t-11),F(t-24),F(t-12),F(t-4),F(t-3),F(t-1)	1,2266	1,0904	45,3
	A(t-5),E(t-5),F(t-23),F(t-12),F(t-5),F(t-3),F(t-1)	1,2505	1,2068	
	A(t-4),E(t-24),F(t-24),F(t-17),F(t-3),F(t-2),F(t-1)	1,2844	1,1021	
	A(t-3),E(t-6),F(t-24),F(t-23),F(t-22),F(t-2),F(t-1)	1,2458	1,1070	
	A(t-5),E(t-8),F(t-24),F(t-22),F(t-13),F(t-4),F(t-1)	1,2146	1,1378	
	A(t-5),E(t-11),F(t-24),F(t-23),F(t-12),F(t-4),F(t-1)	1,2079	1,0435	
2	A(t-4),E(t-2),F(t-16),F(t-8),F(t-4),F(t-2),F(t-1)	1,1439	1,5182	55,3
	A(t-4),E(t-23),F(t-24),F(t-17),F(t-12),F(t-3),F(t-1)	0,9329	1,4300	
	A(t-3),E(t-13),F(t-24),F(t-23),F(t-17),F(t-3),F(t-1)	0,9241	1,3339	
	A(t-4),E(t-19),F(t-24),F(t-23),F(t-17),F(t-3),F(t-1)	0,9393	1,3587	
	A(t-3),E(t-18),F(t-24),F(t-23),F(t-19),F(t-17),F(t-1)	0,9129	1,3447	
	A(t-3),E(t-18),F(t-24),F(t-23),F(t-19),F(t-17),F(t-1)	0,9129	1,3447	

Tabla D.17. PDCP con 5 entradas para la demanda y con una profundidad de 25

Los resultados para una profundidad de 49 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5)$	$E_{(t)}$	$E_{TEST(t)}$	t[<i>min</i>]
1	A(t-5),E(t-3),F(t-24),F(t-23),F(t-13),F(t-3),F(t-1)	1,2538	1,1531	56,0
	A(t-3),E(t),F(t-42),F(t-6),F(t-3),F(t-2),F(t-1)	1,4423	1,4612	
	A(t-3),E(t-21),F(t-43),F(t-39),F(t-24),F(t-3),F(t-1)	1,3367	1,2612	
	A(t-4),E(t-3),F(t-43),F(t-40),F(t-39),F(t-10),F(t-1)	1,3850	1,2405	
	A(t-5),E(t),F(t-43),F(t-42),F(t-34),F(t-24),F(t-17)	1,3551	1,1743	
	A(t-3),E(t-21),F(t-42),F(t-39),F(t-22),F(t-2),F(t-1)	1,3782	1,2211	
2	A(t),E(t-8),F(t-23),F(t-22),F(t-17),F(t-8),F(t-1)	0,9468	1,3408	58,6
	A(t),E(t-8),F(t-44),F(t-23),F(t-16),F(t-8),F(t-1)	1,0044	1,4050	
	A(t-1),E(t-21),F(t-48),F(t-39),F(t-17),F(t-9),F(t-1)	1,0672	1,4709	
	A(t-3),E(t-24),F(t-48),F(t-47),F(t-39),F(t-12),F(t-1)	1,0841	1,4388	
	A(t-4),E(t-24),F(t-47),F(t-39),F(t-36),F(t-35),F(t-1)	1,0659	1,6703	
	A(t-1),E(t-27),F(t-48),F(t-41),F(t-38),F(t-35),F(t-1)	1,0601	1,5227	

Tabla D.18. PDCP con 5 entradas para la demanda y con una profundidad de 49

Los resultados para una profundidad de 73 se muestran en la siguiente tabla.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5)$	$E_{(t)}$	$E_{TEST(t)}$	t[<i>min</i>]
1	A(t-2),E(t),F(t-24),F(t-23),F(t-16),F(t-12),F(t-1)	1,2339	1,0361	58,6
	A(t-4),E(t-2),F(t-69),F(t-23),F(t-17),F(t-16),F(t-1)	1,4310	1,1643	
	A(t-5),E(t),F(t-70),F(t-47),F(t-24),F(t-17),F(t-1)	1,3850	1,1163	
	A(t-5),E(t),F(t-47),F(t-46),F(t-24),F(t-17),F(t-1)	1,3385	1,2023	
	A(t-5),E(t),F(t-72),F(t-69),F(t-49),F(t-42),F(t-23)	1,4215	1,3610	
	A(t-5),E(t-25),F(t-72),F(t-69),F(t-66),F(t-63),F(t-24)	1,5548	1,3020	
2	A(t-4),E(t-3),F(t-24),F(t-23),F(t-17),F(t-3),F(t-1)	0,9325	1,3907	73,7
	A(t-22),E(t-24),F(t-71),F(t-23),F(t-19),F(t-5),F(t-1)	1,0097	1,3529	
	A(t-3),E(t-24),F(t-72),F(t-56),F(t-23),F(t-2),F(t-1)	1,0629	1,4754	
	A(t-3),E(t-35),F(t-72),F(t-71),F(t-56),F(t-19),F(t-1)	1,1175	1,6996	
	A(t-5),E(t-30),F(t-72),F(t-70),F(t-69),F(t-57),F(t-24)	1,1374	1,5419	
	A(t-3),E(t-39),F(t-72),F(t-71),F(t-55),F(t-52),F(t-35)	1,1649	1,6375	

Tabla D.19. PDCP con 5 entradas para la demanda y con una profundidad de 73

La máscara con la que se obtiene el menor error de validación tiene de profundidad 25. Un resumen para las 2 semanas de test se muestra en la Tabla D.20.

sem.	$F(t)=\tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5)$	$E_{(t)}$	\bar{e}_{TE}
1	A(t-5),E(t-11),F(t-24),F(t-23),F(t-12),F(t-4),F(t-1)	1,2079	1,0435
2	A(t-3),E(t-18),F(t-24),F(t-23),F(t-19),F(t-17),F(t-1)	0,9129	1,3447

Tabla D.20. PDCP con 5 entradas para la demanda

D.6. PDCP para la dimensión v_8 - 6 entradas para la variable demanda

Los resultados para una profundidad de 25 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$	$E_{(t)}$	$E_{TEST(t)}$	t[<i>min</i>]
1	A(t-4),E(t-7),F(t-24),F(t-23),F(t-12),F(t-4),F(t-3),F(t-1)	1,1965	1,0495	56,7
	A(t-5),E(t-9),F(t-23),F(t-13),F(t-12),F(t-5),F(t-2),F(t-1)	1,2617	1,2051	
	A(t-5),E(t-11),F(t-24),F(t-23),F(t-12),F(t-4),F(t-3),F(t-1)	1,2063	0,9788	
	A(t-4),E(t-6),F(t-24),F(t-23),F(t-13),F(t-5),F(t-3),F(t-1)	1,1836	1,0887	
	A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,1657	0,9600	
	A(t-5),E(t),F(t-24),F(t-23),F(t-16),F(t-13),F(t-2),F(t-1)	1,2572	1,0387	
2	A(t-3),E(t-6),F(t-23),F(t-22),F(t-21),F(t-20),F(t-18),F(t-4)	1,4012	1,1382	66,5
	A(t-3),E(t-22),F(t-24),F(t-17),F(t-16),F(t-13),F(t-2),F(t-1)	0,9328	1,4594	
	A(t-1),E(t-1),F(t-24),F(t-23),F(t-22),F(t-17),F(t-3),F(t-1)	0,9369	1,3139	
	A(t-3),E(t-2),F(t-24),F(t-23),F(t-21),F(t-16),F(t-12),F(t-1)	0,9250	1,3533	
	A(t-3),E(t-18),F(t-24),F(t-19),F(t-17),F(t-15),F(t-3),F(t-1)	0,9253	1,3888	
	A(t),E(t-13),F(t-24),F(t-23),F(t-21),F(t-17),F(t-16),F(t-1)	0,9122	1,2760	
A(t),E(t-13),F(t-24),F(t-23),F(t-21),F(t-17),F(t-16),F(t-1)	0,9122	1,3076		
A(t),E(t-24),F(t-24),F(t-23),F(t-22),F(t-21),F(t-17),F(t-16)	1,0869	1,4423		

Tabla D.21. PDCP con 6 entradas para la demanda y con una profundidad de 25

Los resultados para una profundidad de 49 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$	$E_{(t)}$	$E_{TEST(t)}$	t[<i>min</i>]
1	A(t-5),E(t),F(t-23),F(t-22),F(t-13),F(t-4),F(t-2),F(t-1)	1,2152	1,2154	66,8
	A(t-4),E(t-2),F(t-23),F(t-17),F(t-11),F(t-3),F(t-2),F(t-1)	1,2464	1,0786	
	A(t-4),E(t-5),F(t-41),F(t-39),F(t-23),F(t-10),F(t-2),F(t-1)	1,3319	1,0440	
	A(t-5),E(t-2),F(t-47),F(t-46),F(t-45),F(t-23),F(t-16),F(t-1)	1,3335	1,1807	
	A(t-5),E(t),F(t-48),F(t-47),F(t-46),F(t-24),F(t-2),F(t-1)	1,3934	1,2236	
	A(t-4),E(t),F(t-23),F(t-17),F(t-8),F(t-4),F(t-2),F(t-1)	1,2732	1,3314	
2	A(t-4),E(t),F(t-23),F(t-17),F(t-8),F(t-4),F(t-2),F(t-1)	1,2732	1,3314	71,1
	A(t-3),E(t-8),F(t-24),F(t-23),F(t-17),F(t-3),F(t-2),F(t-1)	0,9346	1,3638	
	A(t-4),E(t-9),F(t-31),F(t-23),F(t-22),F(t-21),F(t-2),F(t-1)	0,9118	1,3274	
	A(t-4),E(t-9),F(t-48),F(t-39),F(t-17),F(t-12),F(t-2),F(t-1)	1,0045	1,4003	
	A(t-2),E(t-2),F(t-48),F(t-40),F(t-24),F(t-17),F(t-2),F(t-1)	1,0570	1,3279	
	A(t-3),E(t-22),F(t-48),F(t-47),F(t-39),F(t-24),F(t-2),F(t-1)	1,0658	1,3346	
A(t-3),E(t-29),F(t-48),F(t-47),F(t-42),F(t-41),F(t-23),F(t-2)	1,0693	1,4634		
A(t),E(t-7),F(t-48),F(t-40),F(t-34),F(t-32),F(t-25),F(t-12)	1,0718	1,4256		

Tabla D.22. PDCP con 6 entradas para la demanda y con una profundidad de 49

Los resultados para una profundidad de 73 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$	$E_{(t)}$	$E_{TEST(t)}$	t[<i>min</i>]
1	A(t-5),E(t-5),F(t-23),F(t-22),F(t-13),F(t-4),F(t-2),F(t-1)	1,2127	1,2530	79,6
	A(t-5),E(t-16),F(t-42),F(t-24),F(t-23),F(t-16),F(t-4),F(t-1)	1,2680	1,1104	
	A(t-5),E(t-6),F(t-72),F(t-71),F(t-22),F(t-12),F(t-4),F(t-1)	1,5051	1,1866	
	A(t-5),E(t-12),F(t-72),F(t-66),F(t-39),F(t-26),F(t-24),F(t-1)	1,3533	1,2374	
	A(t-3),E(t),F(t-72),F(t-52),F(t-50),F(t-46),F(t-23),F(t-1)	1,3492	1,3541	
	A(t),E(t),F(t-69),F(t-67),F(t-41),F(t-40),F(t-32),F(t-24)	1,4397	1,2184	
2	A(t-5),E(t),F(t-72),F(t-52),F(t-39),F(t-22),F(t-2),F(t-1)	1,3318	1,2555	83,6
	A(t-4),E(t-8),F(t-24),F(t-23),F(t-17),F(t-16),F(t-2),F(t-1)	0,9253	1,3161	
	A(t),E(t-36),F(t-72),F(t-23),F(t-18),F(t-17),F(t-2),F(t-1)	0,9892	1,3367	
	A(t-71),E(t-30),F(t-71),F(t-67),F(t-24),F(t-16),F(t-2),F(t-1)	1,0112	4,5734	
	A(t-57),E(t-25),F(t-72),F(t-71),F(t-19),F(t-16),F(t-2),F(t-1)	1,0635	1,6137	
	A(t-4),E(t-29),F(t-72),F(t-68),F(t-55),F(t-40),F(t-14),F(t-1)	1,0860	1,4935	
A(t-4),E(t-24),F(t-72),F(t-71),F(t-55),F(t-48),F(t-36),F(t-18)	1,1235	1,4521		
A(t),E(t-18),F(t-71),F(t-67),F(t-27),F(t-24),F(t-18),F(t-15)	1,0737	1,3999		

Tabla D.23. PDCP con 6 entradas para la demanda y con una profundidad de 73

La máscara con la que se obtiene el menor error de validación tiene de profundidad 25. Un resumen para las 2 semanas de test se muestra en la Tabla D.24.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$	$E_{(t)}$	$E_{TEST(t)}$
1	A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-16),F(t-2),F(t-1)	1,1657	0,9600
2	A(t),E(t-13),F(t-24),F(t-23),F(t-21),F(t-17),F(t-16),F(t-1)	0,9122	1,2760

Tabla D.24. PDCP con 6 entradas para la demanda

D.7. PDCP para la dimensión v_9 - 7 entradas para la variable demanda

Los resultados para una profundidad de 25 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t),F(t-23),F(t-22),F(t-13),F(t-4),F(t-3),F(t-2),F(t-1)	1,2110	1,2282	69,7
	A(t-5),E(t-17),F(t-24),F(t-17),F(t-12),F(t-4),F(t-3),F(t-2),F(t-1)	1,2046	0,9813	
	A(t-5),E(t-5),F(t-24),F(t-22),F(t-16),F(t-12),F(t-4),F(t-2),F(t-1)	1,1952	1,1161	
	A(t-5),E(t-13),F(t-24),F(t-23),F(t-13),F(t-5),F(t-4),F(t-3),F(t-1)	1,1996	1,0947	
	A(t-3),E(t-1),F(t-24),F(t-23),F(t-22),F(t-16),F(t-13),F(t-2),F(t-1)	1,1817	0,9723	
	A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-16),F(t-12),F(t-2),F(t-1)	1,1818	0,9878	
	A(t-5),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-16),F(t-2),F(t-1)	1,2031	1,0160	
2	A(t-5),E(t-3),F(t-24),F(t-23),F(t-22),F(t-16),F(t-4),F(t-2),F(t-1)	1,1883	1,0234	73,3
	A(t),E(t-19),F(t-24),F(t-13),F(t-11),F(t-9),F(t-3),F(t-2),F(t-1)	1,0331	1,4244	
	A(t-3),E(t-23),F(t-24),F(t-17),F(t-16),F(t-13),F(t-3),F(t-2),F(t-1)	0,9409	1,3282	
	A(t),E(t-22),F(t-24),F(t-23),F(t-18),F(t-16),F(t-12),F(t-2),F(t-1)	0,9548	1,3618	
	A(t-1),E(t-5),F(t-24),F(t-23),F(t-16),F(t-14),F(t-9),F(t-3),F(t-1)	0,9211	1,2981	
	A(t-5),E(t-10),F(t-24),F(t-23),F(t-17),F(t-14),F(t-4),F(t-3),F(t-1)	0,9394	1,3070	
	A(t),E(t-5),F(t-24),F(t-23),F(t-21),F(t-18),F(t-16),F(t-12),F(t-1)	0,9222	1,2970	
A(t),E(t),F(t-24),F(t-23),F(t-20),F(t-18),F(t-17),F(t-15),F(t-1)	0,9298	1,3333		
A(t-1),E(t),F(t-24),F(t-23),F(t-22),F(t-21),F(t-17),F(t-13),F(t-1)	0,9437	1,2957		

Tabla D.25. PDCP con 7 entradas para la demanda y con una profundidad de 25

Los resultados para una profundidad de 49 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-4),E(t-14),F(t-24),F(t-17),F(t-16),F(t-12),F(t-4),F(t-3),F(t-1)	1,2260	1,0116	78,2
	A(t-4),E(t-8),F(t-41),F(t-25),F(t-24),F(t-16),F(t-3),F(t-2),F(t-1)	1,2575	1,1023	
	A(t-3),E(t-5),F(t-41),F(t-32),F(t-24),F(t-22),F(t-14),F(t-2),F(t-1)	1,2971	1,1312	
	A(t-4),E(t-2),F(t-47),F(t-41),F(t-39),F(t-24),F(t-3),F(t-2),F(t-1)	1,2843	1,1573	
	A(t-3),E(t-20),F(t-46),F(t-43),F(t-42),F(t-25),F(t-23),F(t-4),F(t-1)	1,3193	1,2880	
	A(t-4),E(t),F(t-47),F(t-41),F(t-39),F(t-16),F(t-8),F(t-2),F(t-1)	1,2766	1,1176	
	A(t-4),E(t-4),F(t-41),F(t-24),F(t-17),F(t-12),F(t-4),F(t-2),F(t-1)	1,2671	1,1443	
2	A(t-4),E(t),F(t-47),F(t-41),F(t-39),F(t-16),F(t-8),F(t-2),F(t-1)	1,2766	1,1176	75,8
	A(t),E(t-18),F(t-31),F(t-23),F(t-22),F(t-21),F(t-16),F(t-2),F(t-1)	0,9438	1,3649	
	A(t-2),E(t-2),F(t-48),F(t-23),F(t-17),F(t-16),F(t-12),F(t-2),F(t-1)	0,9482	1,3405	
	A(t),E(t-31),F(t-48),F(t-46),F(t-23),F(t-16),F(t-13),F(t-2),F(t-1)	0,9851	1,3879	
	A(t-1),E(t-3),F(t-48),F(t-47),F(t-24),F(t-23),F(t-16),F(t-15),F(t-1)	1,0005	1,2858	
	A(t),E(t-8),F(t-48),F(t-42),F(t-40),F(t-31),F(t-16),F(t-8),F(t-1)	1,0286	1,3843	
	A(t-3),E(t-45),F(t-43),F(t-39),F(t-36),F(t-33),F(t-27),F(t-26),F(t-1)	1,0452	1,4868	
A(t-4),E(t-8),F(t-48),F(t-41),F(t-40),F(t-32),F(t-22),F(t-10),F(t-2)	1,0413	1,3857		
A(t-4),E(t-9),F(t-48),F(t-41),F(t-40),F(t-31),F(t-25),F(t-2),F(t-1)	1,0290	1,4337		

Tabla D.26. PDCP con 7 entradas para la demanda y con una profundidad de 49

Los resultados para una profundidad de 73 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t-7),F(t-24),F(t-22),F(t-16),F(t-12),F(t-4),F(t-2),F(t-1)	1,1726	1,0834	83,9
	A(t-5),E(t-60),F(t-72),F(t-26),F(t-25),F(t-24),F(t-23),F(t-10),F(t-1)	1,2661	1,2422	
	A(t-5),E(t),F(t-68),F(t-67),F(t-23),F(t-18),F(t-3),F(t-2),F(t-1)	1,3829	1,1992	
	A(t),E(t-33),F(t-69),F(t-67),F(t-66),F(t-24),F(t-23),F(t-17),F(t-4)	1,4034	1,1679	
	A(t-2),E(t-3),F(t-72),F(t-50),F(t-46),F(t-24),F(t-23),F(t-4),F(t-1)	1,3014	1,2347	
	A(t-5),E(t),F(t-68),F(t-66),F(t-23),F(t-18),F(t-3),F(t-2),F(t-1)	1,2311	1,2238	
	A(t),E(t-4),F(t-72),F(t-48),F(t-24),F(t-15),F(t-4),F(t-2),F(t-1)	1,1932	1,1567	
	A(t-2),E(t-2),F(t-68),F(t-47),F(t-24),F(t-12),F(t-3),F(t-2),F(t-1)	1,3573	1,1874	
2	A(t-1),E(t-20),F(t-24),F(t-17),F(t-12),F(t-4),F(t-3),F(t-2),F(t-1)	0,9510	1,4137	90,5
	A(t-4),E(t-26),F(t-63),F(t-24),F(t-19),F(t-18),F(t-17),F(t-2),F(t-1)	0,9704	1,5816	
	A(t-3),E(t-29),F(t-72),F(t-57),F(t-27),F(t-24),F(t-20),F(t-2),F(t-1)	1,0004	1,4321	
	A(t-72),E(t-23),F(t-72),F(t-71),F(t-70),F(t-24),F(t-23),F(t-12),F(t-1)	1,0102	4,5852	
	A(t-72),E(t-25),F(t-72),F(t-71),F(t-70),F(t-57),F(t-23),F(t-16),F(t-1)	1,0282	4,4841	
	A(t-2),E(t-17),F(t-72),F(t-71),F(t-56),F(t-47),F(t-16),F(t-13),F(t-1)	1,0911	1,4109	
	A(t-3),E(t-26),F(t-72),F(t-67),F(t-57),F(t-45),F(t-39),F(t-24),F(t-16)	1,0749	1,5111	
	A(t-5),E(t-24),F(t-72),F(t-69),F(t-56),F(t-55),F(t-34),F(t-24),F(t-16)	1,0812	1,4732	

Tabla D.27. PDCP con 7 entradas para la demanda y con una profundidad de 73

La máscara con la que se obtiene el menor error de validación tiene de profundidad 25. Un resumen para las 2 semanas de test se muestra en la Tabla D.28.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$
1	A(t-3),E(t-1),F(t-24),F(t-23),F(t-22),F(t-16),F(t-13),F(t-2),F(t-1)	1,1817	0,9723
2	A(t-1),E(t-5),F(t-24),F(t-23),F(t-16),F(t-14),F(t-9),F(t-3),F(t-1)	0,9211	1,2981

Tabla D.28. PDCP con 7 entradas para la demanda

D.8. PDCP para la dimensión ν_{10} - 8 entradas para la variable demanda

Los resultados para una profundidad de 25 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-4),E(t-16),F(t-24),F(t-17),F(t-12),F(t-5),F(t-4),F(t-3),F(t-2),F(t-1)	1,2357	1,0435	81,0
	A(t-5),E(t-8),F(t-24),F(t-23),F(t-13),F(t-12),F(t-5),F(t-4),F(t-3),F(t-1)	1,1883	1,0194	
	A(t-5),E(t-4),F(t-24),F(t-23),F(t-13),F(t-5),F(t-4),F(t-3),F(t-2),F(t-1)	1,2117	1,1308	
	A(t-5),E(t-3),F(t-24),F(t-23),F(t-22),F(t-16),F(t-4),F(t-3),F(t-2),F(t-1)	1,1874	1,0113	
	A(t-5),E(t-3),F(t-24),F(t-23),F(t-22),F(t-16),F(t-4),F(t-3),F(t-2),F(t-1)	1,1874	1,0113	
	A(t-5),E(t-5),F(t-24),F(t-23),F(t-22),F(t-16),F(t-12),F(t-4),F(t-2),F(t-1)	1,1876	1,0117	
	A(t-2),E(t-3),F(t-24),F(t-23),F(t-22),F(t-17),F(t-16),F(t-14),F(t-2),F(t-1)	1,1924	0,9363	
	A(t-3),E(t-13),F(t-24),F(t-23),F(t-22),F(t-17),F(t-16),F(t-12),F(t-2),F(t-1)	1,2185	0,9724	
	A(t-1),E(t-12),F(t-24),F(t-23),F(t-22),F(t-21),F(t-20),F(t-16),F(t-15),F(t-1)	1,2706	1,0871	
2	A(t-2),E(t-20),F(t-24),F(t-18),F(t-17),F(t-12),F(t-4),F(t-3),F(t-2),F(t-1)	0,9342	1,4315	90,8
	A(t-3),E(t-22),F(t-24),F(t-17),F(t-16),F(t-12),F(t-8),F(t-3),F(t-2),F(t-1)	0,9174	1,3761	
	A(t-4),E(t-23),F(t-24),F(t-17),F(t-16),F(t-12),F(t-8),F(t-3),F(t-2),F(t-1)	0,9268	1,3915	
	A(t),E(t-22),F(t-24),F(t-23),F(t-17),F(t-16),F(t-12),F(t-10),F(t-2),F(t-1)	0,9334	1,3164	
	A(t),E(t-3),F(t-24),F(t-23),F(t-20),F(t-17),F(t-16),F(t-14),F(t-2),F(t-1)	0,9956	1,3571	
	A(t),E(t-4),F(t-24),F(t-23),F(t-22),F(t-17),F(t-16),F(t-8),F(t-2),F(t-1)	0,9046	1,2909	
	A(t-2),E(t-18),F(t-24),F(t-23),F(t-21),F(t-18),F(t-17),F(t-16),F(t-12),F(t-1)	0,9077	1,3302	
	A(t-1),E(t-3),F(t-24),F(t-23),F(t-20),F(t-17),F(t-16),F(t-14),F(t-2),F(t-1)	0,9044	1,2909	
	A(t),E(t-23),F(t-24),F(t-23),F(t-22),F(t-21),F(t-19),F(t-18),F(t-17),F(t-16)	1,0414	1,4419	

Tabla D.29. PDCP con 8 entradas para la demanda y con una profundidad de 25

Los resultados para una profundidad de 49 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8)$	$E_{(t)}$	$E_{TEST(t)}$	t[<i>min</i>]
1	A(t-4),E(t-6),F(t-24),F(t-23),F(t-13),F(t-5),F(t-4),F(t-3),F(t-2),F(t-1)	1,2153	1,1568	71,2
	A(t-5),E(t-10),F(t-42),F(t-24),F(t-23),F(t-22),F(t-11),F(t-3),F(t-2),F(t-1)	1,2466	1,1297	
	A(t-5),E(t),F(t-32),F(t-24),F(t-23),F(t-13),F(t-4),F(t-3),F(t-2),F(t-1)	1,2019	1,1432	
	A(t-5),E(t-2),F(t-47),F(t-46),F(t-25),F(t-24),F(t-23),F(t-9),F(t-2),F(t-1)	1,2795	1,1945	
	A(t-5),E(t-1),F(t-46),F(t-43),F(t-42),F(t-24),F(t-16),F(t-4),F(t-2),F(t-1)	1,2670	1,2200	
	X ⁶			
	X			
2	A(t-2),E(t-30),F(t-24),F(t-18),F(t-17),F(t-16),F(t-12),F(t-3),F(t-2),F(t-1)	0,9211	1,4084	91,7
	A(t),E(t-15),F(t-39),F(t-23),F(t-18),F(t-17),F(t-12),F(t-3),F(t-2),F(t-1)	0,9655	1,4350	
	A(t-1),E(t-3),F(t-48),F(t-24),F(t-23),F(t-17),F(t-15),F(t-9),F(t-3),F(t-1)	0,9563	1,2825	
	A(t-4),E(t-23),F(t-48),F(t-39),F(t-31),F(t-17),F(t-8),F(t-4),F(t-2),F(t-1)	0,9635	1,3838	
	A(t-1),E(t-23),F(t-48),F(t-42),F(t-41),F(t-31),F(t-24),F(t-14),F(t-9),F(t-1)	1,0015	1,3039	
	A(t),E(t-15),F(t-48),F(t-41),F(t-40),F(t-32),F(t-31),F(t-22),F(t-2),F(t-1)	1,0184	1,4133	
	A(t-1),E(t-25),F(t-48),F(t-47),F(t-42),F(t-32),F(t-30),F(t-16),F(t-13),F(t-12)	1,0430	1,4411	
	A(t-2),E(t-31),F(t-48),F(t-47),F(t-46),F(t-45),F(t-42),F(t-41),F(t-16),F(t-1)	1,0705	1,4055	
	A(t),E(t-18),F(t-48),F(t-45),F(t-36),F(t-27),F(t-26),F(t-19),F(t-18),F(t-17)	1,0609	1,5097	

Tabla D.30. PDCP con 8 entradas para la demanda y con una profundidad de 49

Los resultados para una profundidad de 73 se muestran en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8)$	$E_{(t)}$	$E_{TEST(t)}$	t[<i>min</i>]
1	A(t-5),E(t-5),F(t-24),F(t-23),F(t-22),F(t-16),F(t-12),F(t-4),F(t-2),F(t-1)	1,2134	1,0267	75,3
	A(t-5),E(t-18),F(t-66),F(t-24),F(t-22),F(t-16),F(t-5),F(t-3),F(t-2),F(t-1)	1,2529	1,1320	
	A(t-5),E(t),F(t-70),F(t-25),F(t-24),F(t-23),F(t-16),F(t-15),F(t-14),F(t-1)	1,2992	1,1445	
	A(t-2),E(t-12),F(t-69),F(t-50),F(t-47),F(t-25),F(t-24),F(t-23),F(t-4),F(t-1)	1,2623	1,2218	
	A(t-5),E(t-4),F(t-70),F(t-67),F(t-66),F(t-49),F(t-25),F(t-24),F(t-17),F(t-3)	1,3565	1,2123	
	X			
	X			
2	A(t-4),E(t-49),F(t-31),F(t-23),F(t-17),F(t-12),F(t-11),F(t-10),F(t-2),F(t-1)	0,9371	1,4191	95,5
	A(t-1),E(t-18),F(t-31),F(t-24),F(t-23),F(t-21),F(t-16),F(t-15),F(t-2),F(t-1)	0,9116	1,3277	
	A(t-5),E(t-32),F(t-67),F(t-23),F(t-19),F(t-15),F(t-4),F(t-3),F(t-2),F(t-1)	0,9734	1,5751	
	A(t),E(t-28),F(t-72),F(t-67),F(t-55),F(t-24),F(t-14),F(t-12),F(t-2),F(t-1)	0,9874	1,4321	
	A(t),E(t-25),F(t-71),F(t-67),F(t-57),F(t-56),F(t-24),F(t-16),F(t-2),F(t-1)	1,0153	1,5211	
	A(t),E(t-21),F(t-72),F(t-68),F(t-67),F(t-56),F(t-25),F(t-24),F(t-16),F(t-1)	1,0749	1,5577	
	A(t-72),E(t-24),F(t-72),F(t-71),F(t-70),F(t-69),F(t-57),F(t-23),F(t-16),F(t-1)	1,0421	1,4009	
	A(t-5),E(t-18),F(t-72),F(t-67),F(t-57),F(t-39),F(t-33),F(t-32),F(t-24),F(t-18)	1,0047	1,4378	
	A(t-4),E(t-22),F(t-72),F(t-67),F(t-57),F(t-45),F(t-44),F(t-38),F(t-24),F(t-14)	1,0724	1,5089	

Tabla D.31. PDCP con 8 entradas para la demanda y con una profundidad de 73

La máscara con la que se obtiene el menor error de validación tiene de profundidad 25. Un resumen para las 2 semanas de test se muestra en la siguiente tabla.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8)$	$E_{(t)}$	$E_{TEST(t)}$
1	A(t-5),E(t-5),F(t-24),F(t-23),F(t-22),F(t-16),F(t-12),F(t-4),F(t-2),F(t-1)	1,1876	1,0117
2	A(t-1),E(t-3),F(t-24),F(t-23),F(t-20),F(t-17),F(t-16),F(t-14),F(t-2),F(t-1)	0,9044	1,2909

Tabla D.32. PDCP con 8 entradas para la demanda

⁶ En esta condición inicial no se encuentra una solución válida para el periodo de entrenamiento. Se aplica la estructura de vecindad para analizar si existe una solución válida en el entorno de la condición inicial.

E: ANÁLISIS DEL PDCP INCLUYENDO OTRAS TEMPERATURAS

En este anexo se realiza un análisis incluyendo una variable más para la temperatura. En las tablas se detalla tanto el error del grupo de datos de entrenamiento ($E_{(t)}$) como el error del grupo de test ($E_{TEST(t)}$) y el tiempo de simulación en minutos.

E.1. PDCP para la dimensión v_7

Los resultados incluyendo la temperatura mínima de la Sierra se muestran en la Tabla E.1.

sem.	$F(t) = \tilde{f}(a_1, b_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t), B(t-16), E(t), F(t-24), F(t-23), F(t-3), F(t-1)	1,2323	1,2599	50,3
2	A(t-5), B(t), E(t-3), F(t-23), F(t-13), F(t-2), F(t-1)	1,1338	1,4306	55,4

Tabla E.1. PDCP con 4 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Sierra

Los resultados incluyendo la temperatura máxima de la Sierra se muestran en la Tabla E.2.

sem.	$F(t) = \tilde{f}(a_1, c_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t), C(t-24), E(t-7), F(t-24), F(t-3), F(t-2), F(t-1)	1,4148	1,1416	47,8
2	A(t-3), C(t), E(t-10), F(t-21), F(t-18), F(t-17), F(t-1)	0,9326	1,5913	53,9

Tabla E.2. PDCP con 4 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura máxima de la Sierra

Los resultados incluyendo la temperatura mínima de la Costa se muestran en la Tabla E.3.

sem.	$F(t) = \tilde{f}(a_1, d_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-4), D(t-19), E(t-9), F(t-24), F(t-23), F(t-2), F(t-1)	1,1611	1,2034	53,2
2	A(t-3), D(t), E(t-19), F(t-24), F(t-12), F(t-2), F(t-1)	0,9571	1,3478	53,7

Tabla E.3. PDCP con 4 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Costa

E.2. PDCP para la dimensión v_8

Los resultados incluyendo la temperatura mínima de la Sierra se muestran en la Tabla E.4.

sem.	$F(t)=\tilde{f}(a_1,b_1,e_1,f_1,f_2,f_3,f_4,f_5)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-3),B(t-14),E(t-7),F(t-24),F(t-23),F(t-22),F(t-2),F(t-1)	1,2676	1,2389	50,8
2	A(t-5),B(t),E(t-2),F(t-23),F(t-16),F(t-3),F(t-2),F(t-1)	1,1221	1,4576	52,6

Tabla E.4. PDCP con 5 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Sierra

Los resultados incluyendo la temperatura máxima de la Sierra se muestran en la Tabla E.5.

sem.	$F(t)=\tilde{f}(a_1,c_1,e_1,f_1,f_2,f_3,f_4,f_5)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t),C(t-24),E(t),F(t-24),F(t-23),F(t-3),F(t-2),F(t-1)	1,4322	1,3385	55,3
2	A(t),C(t),E(t-14),F(t-24),F(t-23),F(t-21),F(t-18),F(t-1)	0,9164	1,6089	57,3

Tabla E.5. PDCP con 5 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura máxima de la Sierra

Los resultados incluyendo la temperatura mínima de la Costa se muestran en la Tabla E.6.

sem.	$F(t)=\tilde{f}(a_1,d_1,e_1,f_1,f_2,f_3,f_4,f_5)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),D(t-17),E(t-2),F(t-23),F(t-22),F(t-3),F(t-2),F(t-1)	1,1629	1,2802	55,6
2	A(t-2),D(t-1),E(t-18),F(t-24),F(t-23),F(t-18),F(t-16),F(t-1)	0,9480	1,3625	58,7

Tabla E.6. PDCP con 5 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Costa

E.3. PDCP para la dimensión v_9

Los resultados incluyendo la temperatura mínima de la Sierra se muestran en la Tabla E.7.

sem.	$F(t)=\tilde{f}(a_1,b_1,e_1,f_1,f_2,f_3,f_4,f_5,f_6)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-3),B(t-14),E(t-3),F(t-24),F(t-23),F(t-22),F(t-18),F(t-2),F(t-1)	1,2851	1,2966	63,5
2	A(t-5),B(t),E(t),F(t-23),F(t-16),F(t-15),F(t-3),F(t-2),F(t-1)	1,1242	1,5188	66,5

Tabla E.7. PDCP con 6 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Sierra

Los resultados incluyendo la temperatura máxima de la Sierra se muestran en la Tabla E.8.

sem.	$F(t)=\tilde{f}(a_1,c_1,e_1,f_1,f_2,f_3,f_4,f_5,f_6)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t),C(t-24),E(t-11),F(t-24),F(t-23),F(t-22),F(t-16),F(t-15),F(t-1)	1,4168	1,1803	60,9
2	A(t-2),C(t),E(t-15),F(t-24),F(t-23),F(t-20),F(t-18),F(t-17),F(t-1)	0,9248	1,5712	67,1

Tabla E.8. PDCP con 6 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura máxima de la Sierra

Los resultados incluyendo la temperatura mínima de la Costa se muestran en la Tabla E.9.

sem.	$F(t) = \tilde{f}(a_1, d_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-4),D(t-16),E(t-6),F(t-24),F(t-22),F(t-4),F(t-3),F(t-2),F(t-1)	1,1928	1,2696	67,9
2	A(t-2),D(t),E(t-17),F(t-24),F(t-23),F(t-22),F(t-4),F(t-2),F(t-1)	0,9447	1,4311	99,2

Tabla E.9. PDCP con 6 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Costa

E.4. PDCP para la dimensión v_{10}

Los resultados incluyendo la temperatura mínima de la Sierra se muestran en la Tabla E.10.

sem.	$F(t) = \tilde{f}(a_1, b_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-3),B(t-14),E(t-2),F(t-24),F(t-16),F(t-13),F(t-12),F(t-3),F(t-2),F(t-1)	1,2978	1,2377	71,1
2	A(t-5),B(t),E(t-3),F(t-23),F(t-17),F(t-12),F(t-4),F(t-3),F(t-2),F(t-1)	1,1128	1,4999	80,6

Tabla E.10. PDCP con 7 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Sierra

Los resultados incluyendo la temperatura máxima de la Sierra se muestran en la Tabla E.11.

sem.	$F(t) = \tilde{f}(a_1, c_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t),C(t-24),E(t-3),F(t-24),F(t-23),F(t-14),F(t-13),F(t-3),F(t-2),F(t-1)	1,3998	1,1124	99,9
2	A(t-4),C(t-10),E(t-3),F(t-24),F(t-23),F(t-17),F(t-14),F(t-3),F(t-2),F(t-1)	0,9468	1,4000	80,3

Tabla E.11. PDCP con 7 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura máxima de la Sierra

Los resultados incluyendo la temperatura mínima de la Costa se muestran en la Tabla E.12.

sem.	$F(t) = \tilde{f}(a_1, d_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),D(t-18),E(t-5),F(t-24),F(t-23),F(t-17),F(t-4),F(t-3),F(t-2),F(t-1)	1,1909	1,1880	72,2
2	A(t-3),D(t),E(t-18),F(t-24),F(t-23),F(t-20),F(t-18),F(t-17),F(t-2),F(t-1)	0,9483	1,3711	79,5

Tabla E.12. PDCP con 7 entradas para la demanda, 1 entrada para tipo de día, 1 entrada para la temperatura máxima de la Costa e incluyendo 1 entrada para la temperatura mínima de la Costa

F: ANÁLISIS DE SENSIBILIDAD DEL PARÁMETRO “ k_{NN} ”

En este anexo se realiza un análisis de sensibilidad del parámetro k_{NN} . En las tablas se detalla tanto el error del grupo de datos de entrenamiento ($E_{(t)}$) como el error del grupo de test ($E_{TEST(t)}$) y el tiempo de simulación en minutos.

F.1. PDCP para la dimensión v_6 - 4 entradas para la variable demanda

Los resultados para k_{NN} igual a 4 se muestran en la Tabla F.1.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-4),F(t-4),F(t-24),F(t-18),F(t-2),F(t-1)	1,2348	1,1970	24,6
2	A(t),E(t-3),F(t-24),F(t-23),F(t-2),F(t-1)	0,8644	1,2788	28,0

Tabla F.1. PDCP con 4 entradas para la demanda y 4 $_{NN}$

Los resultados para k_{NN} igual a 6 se muestran en la Tabla F.2.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t),F(t-24),F(t-23),F(t-2),F(t-1)	1,3619	1,2188	44,0
2	A(t-3),E(t-18),F(t-24),F(t-18),F(t-3),F(t-1)	0,9926	1,3498	49,1

Tabla F.2. PDCP con 4 entradas para la demanda y 6 $_{NN}$

F.2. PDCP para la dimensión v_7 - 5 entradas para la variable demanda

Los resultados para k_{NN} igual a 4 se muestran en la Tabla F.3.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t),F(t-24),F(t-23),F(t-14),F(t-2),F(t-1)	1,1647	1,0600	38,1
2	A(t-2),E(t-18),F(t-24),F(t-23),F(t-17),F(t-3),F(t-2)	0,8737	1,3557	43,0

Tabla F.3. PDCP con 5 entradas para la demanda y 4 $_{NN}$

Los resultados para k_{NN} igual a 6 se muestran en la Tabla F.4.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t),F(t-24),F(t-23),F(t-4),F(t-2),F(t-1)	1,2663	1,1025	53,1
2	A(t-3),E(t-21),F(t-24),F(t-20),F(t-18),F(t-2),F(t-1)	0,9034	1,3930	62,7

Tabla F.4. PDCP con 5 entradas para la demanda y 6 $_{NN}$

F.3. PDCP para la dimensión v_8 - 6 entradas para la variable demanda

Los resultados para k_{NN} igual a 4 se muestran en la Tabla F.5.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t-5),F(t-24),F(t-23),F(t-18),F(t-3),F(t-2),F(t-1)	1,1615	0,9400	51,1
2	A(t),E(t-22),F(t-24),F(t-17),F(t-16),F(t-12),F(t-2),F(t-1)	0,8785	1,2518	56,3

Tabla F.5. PDCP con 6 entradas para la demanda y 4_{NN}

Los resultados para k_{NN} igual a 6 se muestran en la Tabla F.6.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t-7),F(t-24),F(t-23),F(t-4),F(t-3),F(t-2),F(t-1)	1,1976	1,0778	62,4
2	A(t-3),E(t),F(t-24),F(t-23),F(t-20),F(t-18),F(t-2),F(t-1)	0,9024	1,3432	69,9

Tabla F.6. PDCP con 6 entradas para la demanda y 6_{NN}

F.4. PDCP para la dimensión v_9 - 7 entradas para la variable demanda

Los resultados para k_{NN} igual a 4 se muestran en la Tabla F.7.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t-6),F(t-24),F(t-22),F(t-13),F(t-12),F(t-3),F(t-2),F(t-1)	1,1637	0,9935	63,5
2	A(t),E(t-3),F(t-24),F(t-23),F(t-22),F(t-18),F(t-17),F(t-16),F(t-1)	0,8856	1,2767	68,5

Tabla F.7. PDCP con 7 entradas para la demanda y 4_{NN}

Los resultados para k_{NN} igual a 6 se muestran en la Tabla F.8.

sem.	$F(t) = \tilde{f}(a_1, e_1, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$	$E_{(t)}$	$E_{TEST(t)}$	t[min]
1	A(t-5),E(t-1),F(t-24),F(t-23),F(t-16),F(t-15),F(t-3),F(t-2),F(t-1)	1,2118	1,0165	76,1
2	A(t-2),E(t),F(t-24),F(t-18),F(t-17),F(t-16),F(t-12),F(t-2),F(t-1)	0,9135	1,3300	82,5

Tabla F.8. PDCP con 7 entradas para la demanda y 6_{NN}

G: OFERTA DE GENERACIÓN DEL SISTEMA ECUATORIANO

En este anexo se detalla la oferta de generación que dispone el sistema eléctrico ecuatoriano.


CENTRO NACIONAL DE CONTROL DE ENERGÍA
 DIRECCIÓN DE PLANEAMIENTO
Oferta Hidráulica
PERÍODO: 01 al 31 de mayo del 2006

No.	EMPRESA	CENTRAL	Pot. EFECTIVA (MW)
1	AMBATO	PENÍNSULA	3,0
2	BOLÍVAR	RÍO CHIMBO	1,4
3	CENTRO SUR	SAUCAY	24,0
4	CENTRO SUR	SAYMIRÍN	14,4
5	COTOPAXI	ILLUCHI 1	4,0
6	COTOPAXI	ILLUCHI 2	5,0
7	COTOPAXI	EL ESTADO	1,6
8	E. MUNICIPAL DE AGUA POTABLE	EL CARMEN	8,3
9	E. MUNICIPAL DE AGUA POTABLE	RECUPERADORA	6,0
10	HIDROABANICO	ABANICO	15,4
11	HIDROAGOYÁN	AGOYAN	156,0
12	HIDROAGOYÁN	PUCARÁ	70,0
13	HIDRONACIÓN	M. LANIADO DE WIND	210,0
14	HIDROPAUTE	PAUTE	1.075,0
15	HIDROSIBIMBE	SIBIMBE	15,0
16	QUITO	CUMBAYA	40,0
17	QUITO	NAYON	30,0
18	QUITO	GUANGOPOLO	20,9
19	QUITO	CHILLOS	1,8
20	QUITO	PASOCHOA	4,5
21	QUITO	LORETO	2,0
22	QUITO	PAPALLACTA	3,0
23	REGINAL NORTE	AMBI	8,0
24	REGINAL NORTE	SAN MIGUEL DE CAR	2,9
25	REGINAL NORTE	LA PLAYA	1,2
26	REGINAL NORTE	SAN GABRIEL	0,2
27	REGINAL NORTE	ESPEJO	0,2
28	REGIONAL SUR	CARLOS MORA	2,4
29	RIOBAMBA	ALAO	10,4
30	RIOBAMBA	RÍO BLANCO	3,0



CENTRO NACIONAL DE CONTROL DE ENERGÍA
DIRECCIÓN DE PLANEAMIENTO
Costos Variables de Producción
PERÍODO: 01 al 31 de mayo del 2006

No.	EMPRESA	UNIDAD	TIPO	COSTO ctvs US\$/kWh	Pot. EFECTIVA (MW)
1	TERMOESMERALDAS	ESMERALDAS	BV	2,7038	132,5
2	TERMOPICHINCHA	GUANGOPOLO 1	BD	4,1748	5,2
3	TERMOPICHINCHA	GUANGOPOLO 2	BD	4,2674	5,2
4	TERMOPICHINCHA	GUANGOPOLO 5	BD	4,2728	5,2
5	TERMOPICHINCHA	GUANGOPOLO 6	BD	4,2783	5,2
6	TERMOPICHINCHA	GUANGOPOLO 3	BD	4,2837	5,2
7	TERMOPICHINCHA	GUANGOPOLO 4	BD	4,2892	5,2
8	MACHALA POWER	MACHALA POWER A (*)	G	4,3887	64,9
9	MACHALA POWER	MACHALA POWER B (*)	G	4,4873	65,0
10	ELECAUSTRO	EL DESCANSO 3	BD	4,5902	4,3
11	ELECAUSTRO	EL DESCANSO 4	BD	4,5910	4,3
12	QUITO	G.HERNANDEZ 2	BD	4,5981	5,4
13	ELECTROGUAYAS	TRINITARIA	BV	4,5983	133,0
14	QUITO	G.HERNANDEZ 1	BD	4,6273	5,4
15	QUITO	G.HERNANDEZ 6	BD	4,6273	5,4
16	QUITO	G.HERNANDEZ 4	BD	4,6379	5,4
17	QUITO	G.HERNANDEZ 5	BD	4,6464	5,4
18	QUITO	G.HERNANDEZ 3	BD	4,6829	5,4
19	CEMENTOS SELVA ALEGRE	SELVA ALEGRE	BD	4,7163	13,0
20	ELECAUSTRO	EL DESCANSO 2	BD	4,7753	4,3
21	ELECAUSTRO	EL DESCANSO 1	BD	4,8032	4,3
22	ELECTROGUAYAS	G. ZEVALLOS TV2	BV	5,7125	73,0
23	ELECTROGUAYAS	G. ZEVALLOS TV3	BV	5,7411	73,0
24	CATEG - GENERACIÓN	ANIBAL SANTOS	BV	5,9854	32,5
25	PENINSULA STA. ELENA	LA LIBERTAD 9	D	6,4367	3,0
26	EMELRIOS	CENTRO INDUSTRIAL 4	D	6,4739	2,9
27	PENINSULA STA. ELENA	LA LIBERTAD 10	D	6,5164	2,0
28	MILAGRO	MILAGRO 5	D	6,5551	-
29	REGIONAL MANABI	MIRAFLORES 12	D	6,5774	5,0
30	REGIONAL MANABI	MIRAFLORES 11	D	6,5863	5,0
31	ESMERALDAS	LA PROPICIA 1	D	6,6359	3,6
32	ESMERALDAS	LA PROPICIA 2	D	6,6359	3,6
33	PENINSULA STA. ELENA	LA LIBERTAD 1	D	6,6421	2,0
34	EMELRIOS	CENTRO INDUSTRIAL 3	D	6,6461	2,9
35	CATEG - GENERACIÓN	ALVARO TINAJERO 1	DG	6,6590	46,5
36	EMELRIOS	CENTRO INDUSTRIAL 2	D	6,6684	2,9
37	EMELRIOS	CENTRO INDUSTRIAL 1	D	6,6907	2,9
38	QUITO	LULUNCOTO 12	D	6,7293	2,8
39	ELECTROQUIL	ELECTROQUIL 2	DG	6,8227	46,0
40	PENINSULA STA. ELENA	POSORJA 5	D	6,8415	2,1
41	MILAGRO	MILAGRO 6	D	6,8472	-
42	MILAGRO	MILAGRO 4	D	6,9046	-
43	ELECTROQUIL	ELECTROQUIL 3	DG	6,9486	45,0
44	MILAGRO	MILAGRO 7	D	7,0078	-
45	ELECAUSTRO	MONAY 3	D	7,0310	1,1
46	QUITO	LULUNCOTO 11	D	7,0528	2,7
47	ELECTROQUIL	ELECTROQUIL 1	DG	7,1327	45,0
48	QUITO	LULUNCOTO 13	D	7,1560	2,7
49	AMBATO	BATAN 3	D	7,1911	-
50	ELECTROQUIL	ELECTROQUIL 4	DG	7,2866	45,0
51	ELECAUSTRO	MONAY 2	D	7,2879	1,1
52	REGIONAL EL ORO	ORO CAMBIO 3	D	7,3280	4,4
53	ELECTROGUAYAS	ENRIQUE GARCIA	DG	7,4670	96,0
54	REGIONAL MANABI	MIRAFLORES 10	D	7,5100	2,0
55	REGIONAL MANABI	MIRAFLORES 13	D	7,5100	2,0
56	REGIONAL MANABI	MIRAFLORES 14	D	7,5100	2,0
57	REGIONAL MANABI	MIRAFLORES 15	D	7,5100	2,0
58	REGIONAL MANABI	MIRAFLORES 22	D	7,5249	2,0
59	REGIONAL MANABI	MIRAFLORES 16	D	7,5412	2,0
60	REGIONAL MANABI	MIRAFLORES 18	D	7,5422	2,0
61	ELECAUSTRO	MONAY 1	D	7,5536	1,0
62	REGIONAL MANABI	MIRAFLORES 7	D	7,5776	2,0
63	ULYSSEAS INC.	POWER BARGE I	BV	7,5992	24,0
64	REGIONAL SUR	CATAMAYO 6	D	7,6003	2,5
65	REGIONAL EL ORO	ORO CAMBIO 4	D	7,6245	4,2
66	BOLIVAR	BOLIVAR 1	D	7,6406	1,1
67	REGIONAL MANABI	MIRAFLORES 8	D	7,6647	2,0
68	REGIONAL EL ORO	ORO MACHALA 5	D	7,7896	2,0
69	REGIONAL SUR	CATAMAYO 7	D	7,7955	2,5
70	INTERVISATRADE	VICTORIA II	N	7,8215	105,0
71	CATEG - GENERACIÓN	ALVARO TINAJERO 2	DG	8,0057	35,0
72	PENINSULA STA. ELENA	LA LIBERTAD 11	D	8,0814	2,0
73	REGIONAL SUR	CATAMAYO 2	D	8,0991	1,0
74	REGIONAL EL ORO	ORO MACHALA 4	D	8,1325	2,0
75	RIOBAMBA	RIOBAMBA	D	8,2137	2,0
76	CATEG - GENERACIÓN	ANIBAL SANTOS 1	DG	8,2455	20,0
77	PENINSULA STA. ELENA	PLAYAS 4	D	8,3573	0,7
78	REGIONAL SUR	CATAMAYO 10	D	8,3633	2,2
79	REGIONAL SUR	CATAMAYO 9	D	8,3960	2,2
80	REGIONAL SUR	CATAMAYO 8	D	8,4264	2,2
81	AMBATO	LLIGUA 1	D	8,5977	1,8
82	CATEG - GENERACIÓN	ANIBAL SANTOS 2	DG	8,6685	20,0
83	TERMOPICHINCHA	SANTA ROSA 3	DG	8,7230	17,6
84	CATEG - GENERACIÓN	ANIBAL SANTOS 5	DG	8,7534	18,0
85	CATEG - GENERACIÓN	ANIBAL SANTOS 3	DG	8,7546	14,0
86	TERMOPICHINCHA	SANTA ROSA 2	DG	8,7830	16,9
87	TERMOPICHINCHA	SANTA ROSA 1	DG	8,7854	16,8
88	REGIONAL NORTE	SAN FRANCISCO	D	8,9542	2,0
89	ELECTROGUAYAS	G. ZEVALLOS TG4	DG	8,9900	20,0
90	CATEG - GENERACIÓN	ANIBAL SANTOS 6	DG	9,0192	17,5
91	REGIONAL SUR	CATAMAYO 5	D	9,4048	1,2
92	REGIONAL SUR	CATAMAYO 4	D	9,5036	1,3
93	PENINSULA STA. ELENA	LA LIBERTAD 8	D	9,5351	-
94	REGIONAL MANABI	MIRAFLORES 3	D	9,5733	2,0
95	AMBATO	LLIGUA 2	D	10,0425	1,8
96	ELECAUSTRO	MONAY 5	D	10,2283	1,1
97	ELECAUSTRO	MONAY 4	D	10,2730	1,1
98	ELECAUSTRO	MONAY 6	D	10,6868	1,1

PRECIOS COMBUSTIBLES	SIN TASAS E IMPUESTOS
GALON FUEL OIL 4 (USD)	0,632400
GALON DIESEL 2 (USD)	0,820284
GALON NAFTA (USD)	0,667794
BARRIL RESIDUO SH Guangopolo (USD)	15,134826
BARRIL RESIDUO SH G.Hernández (USD)	14,956746
BARRIL RESIDUO SH S.Alegre (USD)	15,517614
BARRIL RESIDUO Termoesmeraldas (USD)	15,738216
BARRIL RESIDUO ESM 100000 SRW1(USD)	11,220000

OBSERVACIONES

BD Bunker - Diesel
BV Bunker - Vapor
D Diesel
DG Diesel - Gas
G Gas
N Nafta

H: PLANIFICACIÓN DE LA OPERACIÓN

En este anexo se analiza como se realiza la Programación de la Operación en el SSEE del Ecuador [42].

H.1. Planificación Operativa Energética

El CENACE elabora la Planificación Operativa Energética de conformidad a lo establecido en el literal i) del Artículo 24, de la Ley de Régimen del Sector Eléctrico y en el Artículo 6, Capítulo II, del Reglamento de Despacho y Operación del Sistema Nacional Interconectado, siguiendo la metodología que se establece en los Procedimientos de Despacho y Operación.

El objetivo de la Planificación Operativa Energética es establecer una política óptima de la operación de los embalses y uso eficiente de los recursos disponibles de generación, previa la coordinación de mantenimientos efectuada por el CENACE con los generadores y el transmisor, y minimizando los riesgos de falla en el abastecimiento y de vertimientos en los embalses; teniendo en cuenta además, la previsión de las demandas y la aleatoriedad de la oferta y los caudales.

Actualización de la Planificación Operativa Energética

La Planificación Operativa Energética será actualizada cada tres (3) meses, en los meses de enero, abril, julio, octubre, o con una periodicidad menor cuando las circunstancias lo ameriten, para lo cual es necesario incorporar toda la información vigente sobre proyecciones de la demanda, estadística hidrológica, características de las plantas hidráulicas y térmicas, costos variables de producción y topología de la red eléctrica.

El CENACE deberá producir un informe con los resultados obtenidos de la Planificación Operativa Energética, el cual será enviado a los agentes del Mercado Eléctrico Mayorista (MEM). El informe incluirá los siguientes tópicos:

- Datos e hipótesis consideradas.
- Los precios referenciales esperados de generación en barra de mercado estabilizados estacionalmente por banda horaria.
- El valor esperado del agua.

- Disponibilidad de las centrales de generación.
- Programa de mantenimientos.
- Generación prevista por planta hidroeléctrica y por unidad termoeléctrica, para diferentes escenarios hidrológicos representativos.
- Intercambios previstos con otros países, si los hubiere.
- Previsión de la evolución de los niveles de los embalses y de vertimientos, para diferentes escenarios hidrológicos representativos.
- Valor esperado de restricciones al suministro.
- Y otros que se consideren necesarios, a fin de que los agentes tengan la información sobre el posible comportamiento futuro del mercado.

El modelo utilizado para la planificación operativa será aquel que calcule las políticas óptimas de operación de los embalses y las generaciones de las plantas hidroeléctricas y unidades termoeléctricas del Sistema Nacional Interconectado (S.N.I.) que cumplan con la función objetivo de minimizar el valor esperado, llevado a valor presente, de los costos de generación, incluyendo el costo de racionamiento; para lo cual se debe tomar en cuenta lo siguiente:

- Las características operativas de las plantas hidroeléctricas (límites de turbinamiento; para los embalses, se deberá considerar los límites de almacenamiento, los volúmenes de seguridad, los vertimientos, la evaporación, la filtración y las restricciones aguas abajo).
- Las características operativas de las plantas térmicas (capacidad máxima y mínima, restricciones de generación por unidad, costos variables de producción, mantenimientos, consumo de combustible, combustible utilizado, consumo de auxiliares, etc.).
- Aspectos de incertidumbre hidrológica: El modelo permitirá utilizar un modelo estocástico de caudales que represente las características hidrológicas del sistema (estacionalidad, dependencia temporal y espacial de los caudales), secuencias históricas de caudales, o secuencias hidrológicas específicas para condiciones particulares.
- Representación del sistema de transmisión, que tengan en cuenta topología, límites de flujo de potencia en las líneas, pérdidas, restricciones de seguridad, límites de exportación e importación por área eléctrica, etc.
- Representación de la demanda del sistema en etapas semanales o mensuales, conforme a los requerimientos de la planificación operativa energética.

H.2. Objetivos del Planeamiento Eléctrico

El objetivo del planeamiento de la operación eléctrica es garantizar que la operación integrada de los recursos de generación y transmisión cubra la demanda de potencia y energía del Sistema Nacional Interconectado (S.N.I.) cumpliendo con los índices de calidad, seguridad y confiabilidad establecidos en las Regulaciones respectivas. Para ello se utilizarán programas flujos de potencia, cortocircuitos, estabilidad, transitorios electromagnéticos y programas específicos que se requieran de acuerdo a las necesidades.

H.3. Horizontes de Planeamiento Eléctrico de la Operación

El planeamiento eléctrico de la operación se divide en tres etapas: Largo, Mediano y Corto Plazo. En cada una de las etapas se efectúan análisis de estado estacionario, estabilidad transitoria, dinámica y de transitorios electromagnéticos según la profundidad requerida por cada estudio en particular.

Largo Plazo

El objetivo es planear la operación eléctrica del sistema para un horizonte de un año con resolución estacional. La meta será suplir la demanda con los criterios de calidad, seguridad y confiabilidad.

Mediante evaluaciones de estado estacionario, transitorio y dinámico ante fallas, el CENACE hará un informe con periodicidad mínima de un año donde se deberán establecer los siguientes resultados:

- Identificar áreas eléctricas que se caractericen por tener un comportamiento homogéneo o coherente ante fallas.
- Diseño de esquemas de alivio de carga por baja frecuencia y bajo voltaje.
- Determinación de reserva rodante de generación para regulación de frecuencia (primaria: RPF y secundaria: RSF).
- Determinación de reserva fría.
- Voltajes de referencia u objetivo en las principales barras de cada una de las áreas eléctricas.
- Límites de transferencia por líneas o entre áreas eléctricas. Estos límites se determinarán teniendo en cuenta los esquemas y ajustes de las protecciones de los enlaces para evitar los eventos en cascada. Si el aumento de los límites es factible desde el punto de vista de las capacidades del sistema, las protecciones deberán ser ajustadas en consecuencia antes de aumentar los límites.
- Fijar políticas de operación de taps de transformadores.
- Generaciones mínimas de seguridad por área.

- Necesidades de compensación reactiva.
- Recomendaciones para entrada en operación de nuevas instalaciones de transmisión y generación.
- Medidas suplementarias. Las empresas de distribución que deseen cubrirse contra contingencias extremas deberán coordinar con el CENACE la realización de estudios orientados a buscar soluciones para preservar la mayor porción posible del sistema ante la ocurrencia de estos eventos. Después de realizados estos estudios y que el CENACE recomiende las medidas suplementarias del caso, será responsabilidad del agente la implementación de las mismas.

Mediano Plazo

El objetivo es planear la operación eléctrica del sistema para un horizonte de una (1) semana por períodos de demanda (Punta, Media y Base). El objetivo será suplir la demanda con los criterios de calidad, seguridad y confiabilidad establecidos. Mediante evaluaciones de estado estacionario se deberán observar los siguientes resultados:

- Voltajes de referencia u objetivo en las principales barras de cada una de las áreas eléctricas.
- Límites de transferencia por líneas o entre áreas eléctricas, siguiendo los mismos criterios del horizonte de largo plazo.
- Consignas de Operación.
- Análisis de mantenimientos de equipos de generación y transmisión del SNI.
- Estrategias para regulación de voltaje.

Corto Plazo

El objetivo es planear la operación eléctrica del sistema para un horizonte de un (1) día. El objetivo será suplir la demanda con los criterios de calidad, seguridad y confiabilidad establecidos.

Mediante análisis de flujos de potencia se deberán observar los siguientes resultados:

- Estrategias para control de voltaje.
- Recomendar ajustes por razones eléctricas al Despacho Económico horario.
- Análisis de mantenimientos de equipos de generación y transmisión del SNI.

H.4. Programación Semanal

A continuación se detalla la metodología aplicada para la realización de la programación semanal, se establecen los requerimientos de información, descripción del

modelo, plazos de entrega de la información, resultados y se establece las responsabilidades de los agentes del MEM así como del CENACE.

La programación semanal es un proceso de planificación en el corto plazo, abarca un horizonte de 168 horas. La oferta de generación compete en una barra de referencia conocida como “barra de mercado” la cual debe abastecer una demanda. Se considera el sistema de transmisión para calcular las pérdidas y de esta manera encontrar los factores de nodo, se toma en cuenta las restricciones de la red eléctrica. La Programación Semanal (PS) busca la solución de mínimo costo de producción cumpliendo con los objetivos de operación: calidad, economía y seguridad.

La oferta de generación y la demanda del sistema se determina como resultado de la suma de disponibilidades de potencia y energía presentadas por los generadores y de las proyecciones de demanda calculadas por el CENACE o enviadas por los distribuidores, grandes clientes y exportaciones o importaciones para la semana siguiente.

Oferta Hidroeléctrica

Los generadores hidráulicos con capacidad de embalse deben presentar su oferta de generación indicando:

- Programa semanal de mantenimiento con resolución horaria.
- Predicción de caudales en m³/seg. para cada día de la semana.
- Restricciones de aguas abajo, navegación, etc.
- Restricciones operativas: límites mínimo y máximo de generación de las unidades, número de horas mínimo de parada luego de lo cual se podrá ejecutar un nuevo arranque, tiempo de operación mínimo de la unidad, potencia efectiva descontando la reserva.
- Otras restricciones que deban ser tomadas en cuenta y que afecten a la programación semanal.

Los generadores hidráulicos de pasada deben presentar su oferta de generación indicando el programa semanal con resolución horaria de generación, en la cual conste: potencia efectiva y margen de reserva.

Oferta Térmica

Los generadores Térmicos deben presentar su oferta de generación indicando:

- Programa semanal de mantenimiento con resolución horaria.
- Disponibilidad de combustibles en galones.
- Restricciones operativas: Límites mínimo y máximo de generación de las unidades, número de horas mínimo después de parada luego de lo cual se podrá ejecutar un nuevo arranque, tiempo en horas de operación mínimo de la unidad, potencia neta efectiva.

- Otras restricciones que deban ser tomadas en cuenta y que afecten a la programación semanal, como Inflexibilidad de las unidades.

Determinación de la Demanda

El CENACE debe determinar la demanda de los distribuidores, grandes consumidores, así como los exportadores e importadores o en su defecto estos deben enviar las predicciones de sus consumos para un horizonte de 168 periodos (1 semana en adelanto).

Información del Transmisor

El transmisor debe indicar las restricciones del sistema para la semana siguiente enviando la siguiente información:

- Programa semanal con resolución horaria del mantenimiento de líneas de transmisión, transformadores y Subestaciones.
- Restricciones Operativas: Límites de transferencia de las líneas de transmisión, capacidad de transformadores y soporte de reactivos.
- Ingreso en operación de nuevos elementos: Programa semanal con resolución horaria de operación.

Ejecución del Despacho Semanal

El CENACE realizará la programación semanal utilizando un modelo de optimización (Programa de Coordinación hidrotérmica) que represente adecuadamente los recursos energéticos y la red eléctrica del sistema. El modelo realizará la optimización semanal con el criterio de mínimo costo de producción, sujeto a las restricciones operativas del sistema precautelando las condiciones de calidad y seguridad vigentes. En la Figura 5.5 se muestra un diagrama de bloques para el despacho semanal.

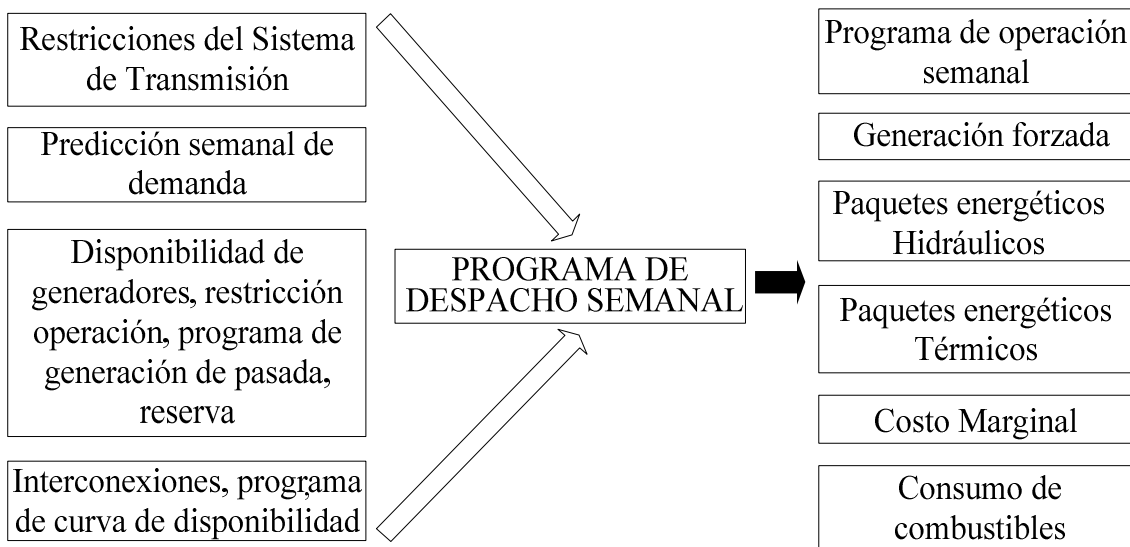


Figura H.1. Programación Semanal en el MEM del Ecuador

La solución, así encontrada deberá ser validada por el área de estudios eléctricos, la misma que determinará la generación forzada necesaria para garantizar condiciones de operación adecuadas así como los márgenes de reserva requeridos para seguridad del sistema. Esta validación se la efectuará mediante simulaciones que representen la red eléctrica. La frecuencia de realización del despacho semanal es de una vez por semana con reajuste diario en caso de haber desviaciones en el transcurso de la semana.

Los Resultados de la Programación semanal incluyen los siguientes aspectos:

- Trayectoria y asignación de cuotas diarias de energía para los embalses.
- Horarios y producción de unidades térmicas.
- Proyección de precios horarios de Mercado ponderados para cada día de la semana por banda horaria.
- Estimación del consumo de combustible.
- Requerimientos de generación forzada.
- Aprobación de los planes de mantenimiento de unidades generadoras y sistema de transmisión.
- Requerimientos de Reserva Adicional de Potencia.
- Reporte a los agentes de todas las restricciones del sistema.
- **Curva de carga horaria del sistema para la semana siguiente.**
- Previsiones de restricciones en el suministro de energía para la siguiente semana.
- Plan de corte por falta de generación o restricciones del sistema.

H.5. Programación Diaria

Despacho Horario

Es el proceso mediante el cual se obtiene para un período de 24 horas el programa horario de generación de los recursos del Mercado Eléctrico Mayorista (MEM) despachados centralmente; y, de las transferencias por las interconexiones internacionales. El Despacho Horario se fundamenta en el criterio de operación a mínimo costo total de producción, en la barra de mercado.

El CENACE tiene a su cargo la responsabilidad de efectuar el cálculo del despacho horario de generación e informarlo a los agentes del MEM.

La información diaria requerida para efectuar el despacho horario es la siguiente:

Demandas

Las empresas distribuidoras y grandes consumidores, deben suministrar diariamente al CENACE la curva de carga diaria de su mercado con resolución horaria, de acuerdo a los formatos establecidos para el efecto. Será responsabilidad del CENACE validar la información recibida, completar los datos faltantes y/o modificar los datos erróneos.

Centrales con Embalse

Las empresas de generación, propietarias de centrales hidráulicas con embalse deben informar diariamente al CENACE, los siguientes datos para cada embalse:

- Nivel del embalse a las 06:00 horas del día en curso, especificando la cota leída en el embalse en metros sobre el nivel del mar (m.s.n.m.) y los millones de metros cúbicos (Mm³) almacenados, correspondientes a esa cota.
- El agua turbinada por la central en el día anterior expresada en millones de m³.
- El agua total vertida por el embalse en el día anterior expresada en millones de m³.
- El promedio de aportes al embalse del día anterior expresados en metros cúbicos por segundo (m³/seg.).
- Las previsiones de caudales de aporte para las próximas veinticuatro horas en metros cúbicos por segundo.
- Mantenimientos, indisponibilidades y restricciones previstas para el día siguiente tanto en el sistema hidráulico asociado al embalse como de las unidades de generación, que puedan afectar la operación de la central. Se debe especificar el período de mantenimiento y las restricciones correspondientes.

Centrales de pasada

Las empresas de generación propietarias de las plantas hidráulicas asociadas a embalses con regulación igual o menor a un día, deben proporcionar diariamente al CENACE la curva de generación horaria prevista para el día siguiente y la correspondiente reserva rodante horaria.

Centrales térmicas e interconexiones internacionales

Las empresas generadoras propietarias de unidades térmicas, deben declarar diariamente al CENACE la disponibilidad de potencia prevista para el día siguiente, a nivel horario para cada unidad generadora o interconexión internacional. Mantenimientos, indisponibilidades y restricciones previstas para el día siguiente que puedan afectar la operación de la unidad. Se debe especificar el período de mantenimiento y las restricciones correspondientes.

Disponibilidad de la red de transmisión

Las empresas transportadoras de energía deben suministrar la siguiente información al CENACE diariamente:

- Mantenimientos de equipos del Sistema de Transmisión, que hayan sido coordinados previamente con los grandes consumidores o con los distribuidores, según sea el caso y por último con el CENACE.
- Deben reportar las indisponibilidades de equipos por mantenimientos correctivos no programados, especificando el período de las desconexiones o el estimativo de la carga a desconectar expresado en valores enteros en MW.
- Disponibilidad de los equipos de compensación inductiva o reactiva.
- Restricciones que deban ser consideradas en el despacho económico diario.

Reserva adicional de potencia

La reserva adicional de potencia determinada en la programación semanal, deberá ser considerada en el cálculo del despacho diario.

En la Figura 5.6 se muestra un diagrama de bloques para la Programación diaria.

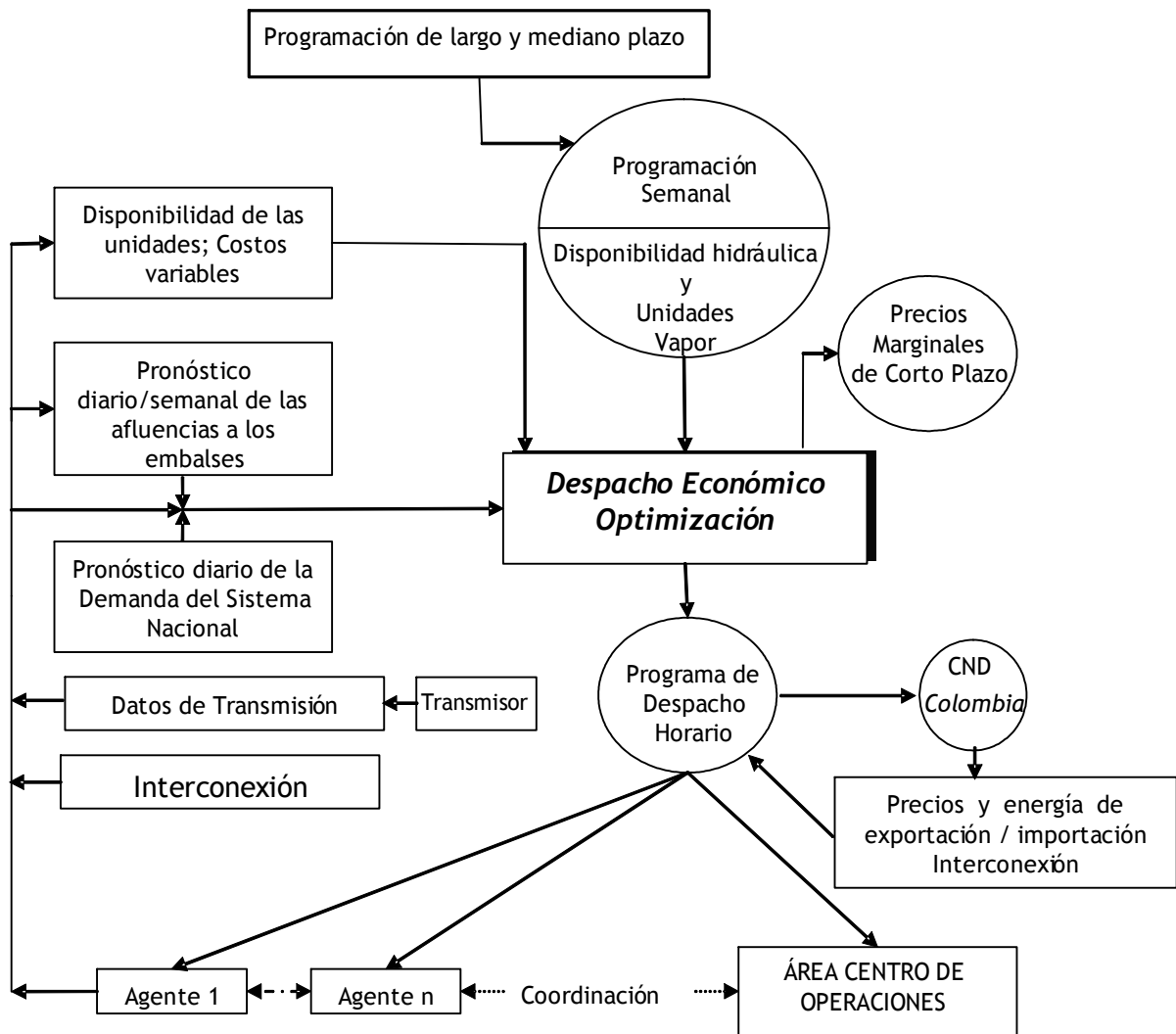


Figura H.2. Programación Diaria en el MEM del Ecuador

H.6. Supervisión y Control de la Operación

La coordinación de la operación del S.N.I. estará a cargo del CENACE. En el sistema de transmisión y los sistemas de distribución, la coordinación de la operación se realizará a través de los respectivos Centros de Operación. El transmisor, el distribuidor y el gran consumidor deberán identificar aquellos equipos de su sistema que tengan asignados técnicamente funciones de transporte, los que estarán sujetos a los mismos derechos y obligaciones que cualquier otro equipamiento del Sistema de Transmisión, en lo que a los aspectos operativos se refiere.

El CENACE realiza la coordinación en tiempo real de la operación de los recursos de generación y transmisión del S.N.I. incluyendo las interconexiones internacionales. Para ello el CENACE coordina la ejecución de las maniobras ya sea con el Centro de Operación del transmisor o de los distribuidores. El CENACE realiza la supervisión del SNI, de los equipos de conexión asociados y de las unidades de generación incluidas en el Despacho Económico. El CENACE hará las funciones de supervisión y coordinación, pero no efectuará comandos directos a los equipos de los Agentes para maniobras en transmisión, distribución, generadores u otros elementos del Sistema, exceptuando los comandos de control de generación (AGC) para los generadores.

Reserva de generación rotante para regulación

Para satisfacer los criterios de calidad y seguridad en la operación del Sistema Nacional Interconectado, ante desbalances entre la oferta y la demanda, el CENACE deberá mantener la reserva de potencia rodante definida en el despacho económico diario programado. Para este propósito, a continuación se señalan algunas instrucciones generales:

- En operación en tiempo real, el CENACE deberá realizar los ajustes necesarios a dicha reserva, para compensar las variaciones entre los valores previstos y los reales, tanto en la oferta como en la **demanda**.
- Todos los generadores en operación, deben cumplir con el porcentaje de reserva requerido para la RPF, establecido en el despacho económico diario programado o redespacho vigente.
- Si en condiciones de operación normal, un generador no cumple con el porcentaje de reserva requerido para la RPF, el CENACE dispondrá al generador se sujete al programa de generación establecido en el despacho económico diario programado o redespacho vigente.
- Si la unidad que participa en la RPF, tiene una disminución en su potencia, el agente generador declarará al CENACE la nueva potencia efectiva máxima, el cual podrá modificar su potencia despachada para mantener el margen de regulación.
- Será obligación de los generadores comunicar al CENACE, cuando una unidad no pueda temporalmente aportar con la RPF, informando las causas de dicha limitación operativa, y el momento a partir del cual volverá a la operación de regulación normal.

- Si una unidad queda imposibilitada de seguir participando en la RPF, el CENACE podrá solicitar despachar la máquina a la potencia efectiva máxima declarada.
- La reserva correspondiente a la RSF se deberá repartir entre todas las unidades habilitadas para tal fin.
- El generador o generadores habilitados para realizar la RSF deberá cumplir con el porcentaje determinado por el CENACE.
- El CENACE vigilará en todo momento, que la diferencia del valor total de la potencia generada y la potencia disponible del grupo de unidades destinadas a la RSF, no viole el porcentaje establecido para la RSF en condiciones de operación normal.

Redespachos

En el artículo 8 del Reglamento de Despacho y Operación del SNI, se establecen las consideraciones que debe tomar el CENACE para el cálculo del despacho económico horario. Así mismo, en el último párrafo de este artículo, se señala que el despacho horario podrá ser modificado durante la ejecución del mismo con el fin de tener en cuenta las condiciones de operación y los recursos del sistema.

Causas de redespacho:

- i) Variación de la disponibilidad de generación. Producto de:
 - Indisponibilidad de alguna de las unidades de generación despachadas.
 - Aumento o disminución en la disponibilidad de unidades de generación.
 - Generación de plantas por pruebas con generación estabilizada que garantice una permanencia superior a una hora.
 - Aumento o disminución de generación
- ii) Variación de los límites en la red de transporte. Cambios topológicos y/o eventos no previstos que impliquen variación en los límites de transferencia de los elementos parte de la red. La generación que ingrese por esta causa será considerada como generación forzada, por lo tanto no afectará en el precio de la energía en el Mercado Ocasional.
- iii) **Variación de la demanda.** Modificación de la demanda en el Mercado, provocando un desvío respecto de la demanda prevista en el despacho vigente en una magnitud que da como resultado el cambio del precio horario de la energía.

Cuando por cambios en el muy corto plazo (menores a 24 horas) de las condiciones climáticas (principalmente temperatura) o cambios considerables en el comportamiento de la demanda debido a factores aleatorios o inesperados, se requiere estimar una nueva proyección de la demanda desde la hora de análisis hasta el final del día de operación de tiempo real. Con esta nueva demanda se realiza un redespacho del parque hidrotérmico y así poder corregir si es del caso la subestimación o la sobreestimación de la demanda.

- iv) Variación de la reserva para regulación de frecuencia. La reserva para regulación de frecuencia se va utilizando en función de las desviaciones que surgen entre la oferta y la demanda. En caso de modificarse la oferta de reserva para regulación de frecuencia ante entradas y/o salidas no previstas de máquinas habilitadas, o inconvenientes informados por el generador que limitan su capacidad de regulación, o modificación en las restricciones previstas de transporte, el CENACE debe realizar un redespacho para recuperar la reserva de potencia para regulación de frecuencia tanto primaria (RPF) como secundaria (RSF).

Como se ve el pronóstico de demanda es una tarea de fundamental importancia para el Operador del Mercado. Por lo que la Dirección de Planeamiento del CENACE requiere:

- Pronóstico de demanda con 1 semana de adelanto es decir necesita el pronóstico de demanda con 168 horas de adelanto (mediano plazo): para la Programación Semanal.
- Pronóstico de demanda con 24 horas de adelanto: para la Programación Diaria.

La Dirección de Operación del CENACE debe realizar la Programación de la Operación de Muy Corto Plazo (de 1 a 24 horas de adelanto), para lo cual requiere:

- Pronóstico de demanda desde 1 a varias horas en adelanto.

REFERENCIAS BIBLIOGRÁFICAS

- 1 Aarts E., “*Simulated Annealing and Boltzman Machines*”, John Willey ans Sons, 1990.
- 2 Argumosa D., “*Planificación de la Expansión de Sistemas de Distribución*”, Instituto de Energía Eléctrica de la Universidad Nacional de San Juan, Argentina, 2002.
- 3 Bezdek J. C., “*Fuzzy Mathematics in Pattern Classification*”, Tesis doctoral, Centro de Matemáticas Aplicadas, Universidad de Cornell, 1973.
- 4 Bhattacharyya G. y Johnson R., “*Statistical concepts and methods*”, John Wiley, Nueva York, 1977.
- 5 Box G. y Jenkins G., “*Time Series Análisis: forecasting and control*”, Prentice Hall, Englewood Cliffs - NJ, 1994.
- 6 Carvajal R. y Nebot A., “*Growth model for white shrimp in semi-intensive farming using Inductive Reasoning methodology*”, Computers and Electronics in Agriculture, 19 (2), pp 187-210, 1998.
- 7 Cellier F., “*Qualitative Modeling and Simulation – Promise or Illusion*”, Proceedings Winter Simulation Conference, 1991.
- 8 Cellier F., “*Continuous System Modeling*”, Springer - Verlag, Nueva York - USA, 1991.
- 9 Cellier F., “*Mixed Quantitative and Qualitative Modeling: Means for Dealing with System Uncertainty*”, Proceedings 15th Benelex Meeting on Systems and Control, Holanda, 1996.
- 10 Cellier F., Nebot A. Mugica F. y De Albornoz A., “*Combined Qualitative/Quantitative Models of Continuos-Time Processes using Fuzzy Inductive Reasoning Techniques*”, International Journal of General Systems, 1996.
- 11 Cellier F. y Yandell D., “*Saps II: A new implementation of the system approach problem solver*”, International Journal of General Systems, 1987, 307-322.
- 12 Chen C. S., Tzeng Y. M. Y Hwang J. C., “*The Application of Artificial Neural Networks to Substation Load Forecasting*”, Electric Power Systems Research 38, 1996.
- 13 Chiu S. L., “*Fuzzy model identification based on cluster estimation*”, Journal of Intelligent and Fuzzy Systems, 2 (3), 1994.

- 14 Dasgupta, D. y Michalewicz Z., “*Evolutionary Algorithms in engineering applications*”, Springer, 1997.
- 15 De Albornoz, A., “*Inductive Reasoning and Reconstruction Analysis: Two Complementary tools for Qualitative Fault Monitoring and Decision Support in Large-Scale Systems*”, Tesis doctoral, Universidad Politécnica de Cataluña, España, 1996.
- 16 Doña V., “*Utilización óptima temporal del criterio de seguridad (n-1) en sistemas eléctricos interconectados débilmente mallados*”, Tesis doctoral, Universidad Nacional de San Juan, Argentina, 1996.
- 17 Goldberg D., “*Genetic Algorithms in Search, Optimization and Machine Learning*”, Addison-Wesley, 1989.
- 18 Götze U. y Bloech J., “*Investitionsrechnung: Modelle und Analysen zur Beurteilung von Investitionsvorhaben*”, Springer Verlag, 1993.
- 19 Hagan M. T. y Behr S. M., “*The Times Series Approach To Short Term Load Forecasting*”, IEEE Transactions on Power System, Vol. PWRS-2, No. 3, 1987.
- 20 Haykin S., “*Neural Networks: A Comprehensive Foundation*”, Prentice Hall Inc., USA, 1999.
- 21 Hinojosa V. “*Pronóstico de Demanda de Corto Plazo en Sistemas de Suministro de Energía Eléctrica utilizando Redes Recurrentes (Jordan)*”, XI ERIAC, Paraguay, 2005.
- 22 Hoese A., “*Modelling, simulation and analysis of renewable electrical energy generation systems*”, Tesis de Maestría, Universidad de Kaiserslautern, Alemania, .1993.
- 23 Hoese A., “*A new algorithm of combinatorial optimization applied to the evaluation of a stand-alone electric generation system*”, Proceedings of the 24th Argentine Conference on Informatics and Operations Research, Argentina, pp. 47-54, 1995.
- 24 Hoese A., “*Planificación Óptima de la Expansión de Sistemas aislados con participación de Energías primarias renovables*”, Tesis doctoral, Universidad Nacional de San Juan, Argentina, 2000.
- 25 Instituto de Energía Eléctrica, “*Presentación General del Problema de Optimización de Sistemas de Suministro de Energía Eléctrica*”, Universidad Nacional de San Juan, Argentina, 2000.
- 26 Jang J-S. R., Sun C-T., y Mizutani E. “*Neuro-Fuzzy and Soft Computing*”, Prentice Hall, New Jersey, 1997.
- 27 Klir G. J., “*Architecture of System Problem Solving*”, Plenum Press, New York, 1985.
- 28 Klir G. J., “*Inductive Systems Modelling: An Overview, Modelling and Simulation Methodology: Knowledge Systems’ Paradigms*”, M. S. Elzas, T. I. Ören and B. P. Zeigler eds., Elsevier Science Publishers B. V, 1989.
- 29 Li D. y Cellier F., “*Fuzzy Measures in Inductive Reasoning, In Proceedings of the Winter Simulation Conference*”, Los Angeles, 1990.

- 30 López J., "*Time Series Prediction Using Inductive Reasoning Techniques*", Tesis doctoral, Universidad Politécnica de Cataluña, España, 1999.
- 31 Lotufo A. D. y Minussi C. R., "*Electric Power Systems Load Forecasting: A Survey*", IEEE Power Tech, Hungría, 1999.
- 32 Makridakis S. y Wheelwright S., "*Forecasting: Methods and Applications*", John Wiley & Sons Inc., USA, 1978.
- 33 Miranda V., Srinivasan D., y Proenca L. "*Evolutionary computation in power systems*", Electric Power and Energy Systems, Vol. 20, pp. 89-98, 1998.
- 34 Mirats J. M., "*Qualitative Modelling of Complex Systems by means of Fuzzy Inductive Reasoning. Variable Selection and Search Space Reduction*", Tesis doctoral, Universidad Politécnica de Cataluña, España, 2001.
- 35 Moller M. F., "*A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning*", Neural Networks, vol. 6, pp. 525-533, 1993.
- 36 Moorthy M., "*Mixed Structural and Behavioral Models for Predicting the Future Behavior of Some Aspects of the Macroeconomy*", Tesis de maestría, Universidad de Arizona, USA, 1999.
- 37 Mugica, F., "*Diseño Sistemático de Controladores Difusos usando Razonamiento Inductivo Fuzzy*", Tesis doctoral, Universidad Politécnica de Cataluña, España, 1995.
- 38 Mugica F. y Cellier F., "*A New Fuzzy Inferencing Method for Inductive Reasoning*", Sixth International Symposium on Artificial Intelligence, México, 1993.
- 39 Nebot, A., "*Qualitative Modeling and Simulation of Biomedical Systems using Fuzzy Inductive Reasoning*", Tesis doctoral, Universidad Politécnica de Cataluña, España, 1994.
- 40 Nebot, A. y Mugica F., "*Using causal relevancy for the selection of models within FIR qualitative modelling and simulation environment*", Advances in Artificial Intelligence and Engineering Cybernetics, Vol. III, pp 137-142, 1994.
- 41 Nguyen D. y Widrow B., "*Improving the Learning Speed of 2-layer Neural Network by Choosing Initial Values of the Adaptive Weights*", Proceedings of the International Joint Conference on Neural Networks, vol. 3, pp. 21-26, 1990.
- 42 Regulación CONELEC No. 006/00, "*Procedimientos de Despacho y Operación*", Resolución No. 0125/00, 2000.
- 43 Samaniego A., "*Pronóstico de demanda de corto plazo del Sistema Nacional Interconectado mediante Modelos ARMA*", Tesis de diplomado, Escuela Politécnica Nacional, Ecuador, 2002.
- 44 Sánchez G., "*Inteligencia artificial aplicada al diseño de sistemas de generación de energía eléctrica*", Tesis de diplomado, Universidad Nacional de San Juan, Argentina, 1998.
- 45 Sánchez G. y Hoese A., "*Comparing evolutionary algorithms applied to the optimal design of generation power systems*", Argentine Symposium on Artificial Intelligence (ASAI'99), Buenos Aires, Argentina, pp. 297-305, 1999.

- 46 Taylor J. R., “*An Introduction to Error Analysis*”, University Science Books, USA, 1997.
- 47 Uyttenhove H., “*SAPS – System Approach Problem Solver*”, Tesis doctoral, SUNY Binghamton, USA, 1979.
- 48 Van Welden D., A., “*Induction of Predictive Models for Dynamical Systems via Data Mining*”, Tesis doctoral, Universidad de Gent, Alemania, 2000.
- 49 Willis H. L., “*Spatial Electric Load Forecasting*”, ABB Systems Control Division Cary, North Carolina, 1996.
- 50 Yang H. T. y Huang C. M., “*A New Short Term Load Forecasting Approach Using Self-Organizing Fuzzy ARMAX Models*”, IEEE Transactions on Power System, Vol. 13, No. 1, 1998.

Curriculum Vitae

Nombre: Víctor Hugo Hinojosa Mateus.

Fecha de Nacimiento: 30 de enero de 1975 en Quito, Ecuador.

Padres: Víctor Hugo y Luz Margarita Piedad.

Estado civil: casado desde el 12 de agosto de 2000 con Viviana Tapia.

Hijos: Ariana Monserrat y Pía Valentina.

Escuela: 1980-1986 Escuela Luís Felipe Borja – Machachi
1986-1992 Colegio La Salle – Conocoto.
Egresó: julio de 1991 – Bachiller en Ciencias.

Estudios de Grado: 1992-1999 Ingeniería Eléctrica en la Escuela Politécnica Nacional, Quito – Ecuador.

Becas obtenidas: desde el 01/08/2001 hasta el 31/07/2005 Becario del Servicio Alemán de Intercambio Académico (DAAD).

Distinciones: 1986 – primer escolta de la bandera tricolor “Escuela Luís Felipe Borja”.

Actividad laboral: Desde febrero de 2006 se desempeña como Investigador en el Área de Investigación y Desarrollo del Centro Nacional de Control de Energía – CENACE del Ecuador.
Desde marzo de 2006 colabora con cátedras en la Facultad de Ingeniería Eléctrica de la Escuela Politécnica Nacional – Ecuador.
